

SURVEY ARTICLE

- Psychological and Social Problems of Man in Space—a Literature Survey B. D. Goodman 863

CONTRIBUTED ARTICLES

- Feasibility of Turbulent Vortex Containment in the Gaseous Fission Rocket M. L. Rosenzweig, W. S. Lewellen and J. L. Kerrebrock 873
- Propulsion System Using a Cavity Reactor and Magnetohydrodynamic Generator Richard J. Rosa 884
- Photochemistry and Space Power Generation J. N. Pitts Jr., J. David Margerum and Wm. E. McKee 890
- Internal Design Consideration for Cavity-Type Solar Absorbers Charles W. Stephens and Alan M. Haire 896
- High Temperature Rarefied Hypersonic Flow Over a Flat Plate H. T. Nagamatsu, R. E. Sheer Jr. and J. R. Schmid 902
- Analysis of Convective Heat Transfer in Rocket Nozzles Ernest Mayer 911
- Ignition Characteristics of Metals and Alloys L. E. Dean and W. R. Thompson 917
- The Artificial Earth Satellite—a New Geodetic Tool Bruce C. Murray 924
- Relating Geodetic Latitude and Altitude to Geocentric Latitude and Radius Vector E. W. Purcell and W. B. Cowan 932
- Rocket Boost Vehicle Mission Optimizations R. T. Stancil and L. J. Kulakowski 935
- Optimum Maneuvers for Launching Satellites into Circular Orbits of Arbitrary Radius and Inclination J. P. Carotens and T. N. Edinburg 943
- Self-Contained Navigational System for Determination of Orbital Elements of a Satellite K. N. Satyendra and R. E. Bradford 949

TECHNICAL NOTES

- Surface Temperatures of High-Speed, Radiation Cooled Bodies in Dissociating Atmospheres Daniel E. Rosner 1013
- Satellite Life Duration Thomas L. Vincent 1015
- High-Energy, Low-Thrust Jupiter Missions Robert H. Fox 1016
- Maximum Error in Total Radiosity Measurements Due to Non-Grayness of Samples D. K. Edwards and E. E. Nelson 1021
- Hypersonic Viscous Effects in Wind Tunnels Robert H. Johnson 1022
- On "Numerical Comparison Between Exact and Approximate Theories of Hypersonic Inviscid Flow Past Slender Blunt-Nosed Bodies" K. H. Cheng and A. L. Chang 1024
- Effect of Acoustic Environment on the Burning Rate of Double-Burn Solid Propellants J. E. Crump and E. W. Price 1028
- Advanced Ignition System for Solid Propellant Rocket Motors J. J. Pripi 1029
- Secular Gravitational Torque on a Satellite in a Circular Orbit Russell A. Nidey 1032

DEPARTMENTS

- Patents 1033
- Book Notes 1034
- Technical Literature Digest 1035

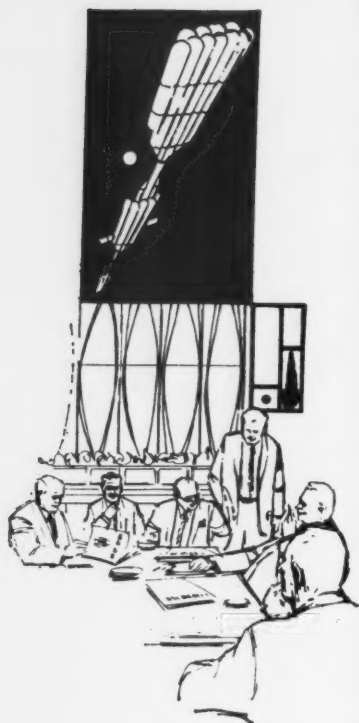
New York Coliseum

October 9-15, 1961

AMERICAN ROCKET SOCIETY

SPACE FLIGHT REPORT

TO THE NATION



Through the medium of the ARS SPACE FLIGHT REPORT TO THE NATION, the American Rocket Society, during the week of October 9-15, 1961, in New York Coliseum, will present a thorough and comprehensive review of the U. S. space program.

The SPACE FLIGHT REPORT TO THE NATION will present for the first time under one roof a complete review of the most significant work being done in each of the major technical disciplines contributing to current progress in rocketry, missiles, and space flight, and one of the largest and most extensive technical exhibits ever undertaken in these fields.

More than 300 technical papers will detail major space projects; will explore problems that must be solved to insure further, faster progress; and will stimulate penetrating discussions by outstanding authorities in the audience.

There will be three floors of technical exhibits and displays by more than 350 manufacturers, suppliers, engineering organizations, and governmental agencies, disclosing latest achievements, concepts, products, and ideas.

AR
RE
NA

• Un
place
all th
U. S.

• Fo
sions,
able l
for in

• Th
panels
space
global
techni

• A
"Exhi
the fi
with
the bo

• Ma
DOD,
bility,
area a

• The
ARS
design



For fur
program
contact

A
REP

Am

50

ARS SPACE FLIGHT REPORT TO THE NATION will present:

- Under one roof, at one time, in one place the most comprehensive review of all the contributing technologies to the U. S. space effort.

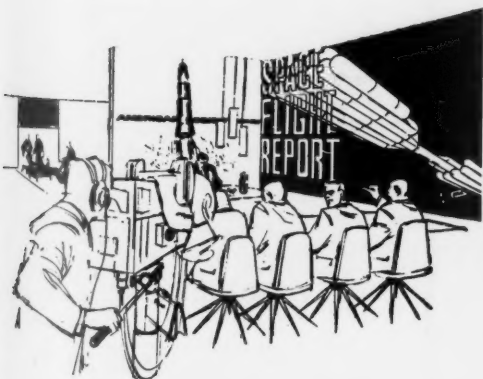
- Forty-two specialized technical sessions, all on the same floor with comfortable lounge areas and conference rooms for informal discussions.

- Three unique Space Flight Report panels that will explore and interpret space missions, space vehicles, and the global impact of space events for the technical community at large.

- A new exhibiting technique called "Exhibiting in Depth", permitting for the first time 3-dimensional displays with the main aisle running *through* the booths instead of by them.

- Main exhibit organized by NASA and DOD, representing U. S. space capability, which will be 88'x50' in floor area and will extend 60' into the air.

- The history, purposes, and plans of ARS graphically displayed in a booth designed by Raymond Loewy.



For further information about reservations, program information or exhibit space contact

ARS SPACE FLIGHT REPORT TO THE NATION

American Rocket Society

500 Fifth Ave., New York 36, N. Y.

Honorary Committee

Wernher von Braun, Chairman

Harold Boyer

Detlev W. Bronk

Overton Brooks

Milton U. Clauser

James R. Dempsey

Gen. James H. Doolittle

C. Stark Draper

Hugh L. Dryden

Trevor Gardner

Mrs. Robert H. Goddard

Harry F. Guggenheim

Willis M. Hawkins

Rear Adm. John T. Hayward

Samuel K. Hoffman

Robert H. Jewett

Arthur Kantrowitz

Dan A. Kimball

Maj. Gen. John B. Medaris

George F. Metcalf

Maj. Gen. Don R. Ostrander

William H. Pickering

Richard W. Porter

Simon Ramo

Harold W. Ritchey

Lt. Gen. Bernard A. Schriever

Howard S. Seifert

Abe Silverstein

Col. John P. Stapp

George P. Sutton

Maj. Gen. Holger N. Toftoy

George S. Trimble, Jr.

Lt. Gen. Arthur G. Trudeau

Harold Urey

Theodore von Karman

Alan Waterman

Elmer P. Wheaton

Herbert F. York

ARS JOURNAL

A PUBLICATION OF THE AMERICAN ROCKET SOCIETY

EDITOR Martin Summerfield
ASSOCIATE TECHNICAL EDITOR Irvin Glassman
MANAGING EDITOR Margaret Sherman
STAFF EDITOR Barbara Nowak
ART EDITOR John Culin

ASSOCIATE EDITORS

Igor Jurkevich, G. E. Space Sciences Laboratory, Russian Supplement;
 George F. McLaughlin, Patents; Charles J. Mundo Jr., Raytheon Co.,
 Guidance; Bernard H. Paiewonsky, Aeronautical Research Associates
 of Princeton, Flight Mechanics; M. H. Smith, Princeton University,
 Technical Literature Digest

ASSISTANT EDITORS

Carol Falcetti, Eleanor Rosman, Carol Rubenstein

DIRECTOR OF MARKETING

Owen A. Kean

ADVERTISING PRODUCTION MANAGER

Walter Brunke

ADVERTISING REPRESENTATIVES

New York

D. C. Emery and Associates
 400 Madison Ave., New York, N. Y.
 Telephone: Plaza 9-7460

Chicago

Jim Summers and Associates
 35 E. Wacker Dr., Chicago, Ill.
 Telephone: Andover 3-1154

Atlanta

Joe H. Howell
 1776 Peachtree Bldg., Atlanta 9, Ga.
 Telephone: 873-2136

Los Angeles

James C. Galloway and Co.
 6535 Wilshire Blvd., Los Angeles, Calif.
 Telephone: Olive 3-3223

Detroit

R. F. Pickrell and Associates
 318 Stephenson Bldg., Detroit, Mich.
 Telephone: Trinity 1-0790

London

B. C. Nichols
 151 Fleet St.
 London E. C. 4, England

American Rocket Society

500 Fifth Avenue, New York 36, N. Y.

Founded 1930

OFFICERS

President	Harold W. Ritchey
Vice-President	William H. Pickering
Executive Secretary	James J. Harford
Treasurer	Robert M. Lawrence
General Counsel	Andrew G. Haley
Director of Publications	Irwin Hersey

BOARD OF DIRECTORS

(Terms expire on dates indicated)

Ali B. Cambel	1962	Samuel Herrick	1963
Richard B. Canright	1962	Arthur Kantrowitz	1963
William J. Cecka, Jr.	1963	A. K. Oppenheim	1961
James R. Dempsey	1961	Simon Ramo	1963
Herbert Friedman	1962	David G. Simons	1961
George Gerard	1961	John L. Sloop	1961
Robert A. Gross	1962	Martin Summerfield	1962
		Abe M. Zarem	1963

TECHNICAL COMMITTEES AND CHAIRMEN

Aerodynamics , Robert M. L. Baker, Jr.	Physics of the Atmosphere and Space , Herbert Friedman
Communications and Instrumentation , Frank W. Lehan	Power Systems , John H. Huth
Electric Propulsion , Ernst Stuhlinger	Propellants and Combustion , Peter L. Nichols, Jr.
Guidance and Control , James S. Farrior	Ramjets , William H. Avery
Human Factors and Bioastronautics , Eugene E. Konecki	Solid Propellant Rockets , G. Daniel Brewer
Hypersonics , William H. Dorrance	Space Law and Sociology , Andrew G. Haley
Liquid Rockets , Martin Goldsmith	Structures and Materials , George Gerard
Magnetohydrodynamics , Milton M. Slawsky	Test, Operations, and Support , Bernhard L. Dorman
Missiles and Space Vehicles , William M. Duke	Underwater Propulsion , Herman E. Sheets
Nuclear Propulsion , Robert W. Bussard	

Scope of ARS JOURNAL

This Journal is devoted to the advancement of astronautics through the dissemination of original papers disclosing new scientific knowledge and basic applications of such knowledge. The sciences of astronautics are understood here to embrace selected aspects of jet and rocket propulsion, spaceflight mechanics, high speed aerodynamics, flight guidance, space communications, atmospheric and outer space physics, materials and structures, human engineering, overall system analysis and possibly certain other scientific areas. The selection of papers to be printed will be governed by the pertinence of the topic to the field of astronautics, by the current or probable future significance of the research, and by the importance of distributing the information to the members of the Society and to the profession at large.

Information for Authors

Manuscript must be as brief as the proper presentation of the ideas will allow. Exclusion of dispensable material and conciseness of expression will influence the Editors' acceptance of a manuscript. In terms of standard-size double-spaced typed pages, a typical maximum length is 22 pages of text (including equations), 1 page of references, 1 page of abstract and 12 illustrations. Fewer illustrations permit more text, and vice versa. Greater length will be acceptable only in exceptional cases.

Short manuscripts, not more than one quarter of the maximum length stated for full articles, may qualify for publication as Technical Notes or Technical Comments. They may be devoted to new developments requiring prompt disclosure or to comments on previously published papers. Such manuscripts are published within a few months of the date of receipt.

Sponsored manuscripts are published occasionally as an ARS service to the industry. A manuscript that does not qualify for publication, according to the above-stated requirements as to subject, scope or length, but which nevertheless deserves widespread distribution among jet propulsion engineers, may be printed as an extra part of the Journal or as a special supplement, if the author or his sponsor will reimburse the Society for actual publication costs. Estimates are available on request. Acknowledgment of such financial sponsorship appears as a footnote on the first page of the article. Publication is prompt since such papers are not in the ordinary backlog.

Manuscripts must be double spaced on one side of paper only with wide margins to allow for instructions to printer. Include a 100 to 200 word abstract. State the authors' positions and affiliations in a footnote on the first page. Equations and symbols may be handwritten or typewritten; clarity for the printer is essential. Greek letters and unusual symbols should be identified in the margin. If handwritten, distinguish between capital and lower case letters, and indicate subscripts and superscripts. References are to be grouped at the end of the manuscript and are to be given as follows. For journal articles: Authors first, then title, journal, volume, year, page numbers; for books: Authors first, then title, publisher, city, edition and page or chapter numbers. Line drawings must be clear and sharp to make clear engravings. Use black ink on white paper or tracing cloth. Lettering should be large enough to be legible after reduction. Photographs should be glossy prints, not matte or semi-matte. Each illustration must have a legend; legends should be listed in order on a separate sheet.

Manuscripts must be accompanied by written assurance as to security clearance in the event the subject matter lies in a classified area or if the paper originates under government sponsorship. Full responsibility rests with the author.

Preprints of papers presented at ARS meetings are automatically considered for publication.

Submit manuscripts in duplicate (original plus first carbon, with two sets of illustrations) to the Managing Editor, ARS JOURNAL, 500 Fifth Avenue, New York 36, N.Y.

ARS JOURNAL is published monthly by the American Rocket Society, Inc. and the American Interplanetary Society at 20th & Northampton Sts., Easton, Pa., U. S. A. Editorial offices: 500 Fifth Ave., New York 36, N. Y. Price: \$18.00 per year, \$3.00 per single copy. Second-class postage paid at Easton, Pa., with additional entry at New York, N. Y. This publication is authorized to be mailed at the special rates of postage prescribed by Section 132.122. Notice of change of address should be sent to the Secretary, ARS, at least 30 days prior to publication. Opinions expressed herein are the authors' and do not necessarily reflect the views of the editors or of the Society. © Copyright 1961 by the American Rocket Society, Inc.

When, as was commonly the case, I had none to commune with, I used to raise the echoes by striking with a paddle on the side of my boat, filling the surrounding woods with circling and dilating sounds.

—HENRY DAVID THOREAU, *Walden*

Psychological and Social Problems of Man in Space—a Literature Survey

B. D. GOODMAN¹

System Development Corp.
Santa Monica, Calif.

What type of man will be able to endure for months or even years the vast silence and loneliness of space, far removed from the sounds and sights of his natural environment? What type of man can remain alert and maintain his performance, deprived of ordinary sensory stimuli, enclosed in the cramped quarters of a space capsule as it leaves Earth and all that is familiar?

It is the purpose of this bibliography to bring together the reports, books and periodical articles published through the early part of 1961 dealing with the specific area of behavioral science related to space flight, or as it is sometimes called "space psychology." This area includes problems of confinement, isolation, sensory deprivation, weightlessness, psychological assessment and training, motivation and morale, emotional stability, boredom and fatigue, performance under stress, and work load.

A REVIEW of space literature shows that the primary emphasis in research has been on engineering—designing the space vehicle to get man into space, and providing the proper closed ecological system, the necessary controls, displays and equipment to make it possible for him to survive there. This phase of research has involved studying man and his physiological needs, and analyzing human tolerances to environmental variances. An accompanying, but less emphasized phase has been concerned with the study of the psychological and social problems of man in space.

Such problems warrant careful study for prolonged orbits or interplanetary flights lasting months or even years. It is not to be expected that they would create hazards in a short-term space flight such as one orbit around Earth.

An alphabetized and annotated bibliography will be given of the space literature pertaining to the psychological and social problems of man in space. Citations listed are unclassified unless otherwise noted. All titles are unclassified. To facilitate ordering items listed in the Technical Abstract Bulletin (TAB) of the Armed Services Technical Information Agency, ASTIA document (AD) numbers have been given when known.

General Surveys

The complexities of the psychophysiological (human) factors involved in sending man out into space have been examined in numerous general surveys.

In an excellent historical account, Generales (49)² indicates the wide scope of research in space medicine. As early as 1949, Armstrong, Haber and Strughold (9) included psychological factors in an analysis of the aeromedical problems of space travel.

Konecni (86) presents a clear-cut tabular breakdown of possible human factor problem areas under main subject headings of physiological, human engineering, and psychological-social. A U. S. Senate report (146) defines the broad field of investigation in life sciences related to the national space effort; Flickinger (43 and 44) reviews the U. S. Air Force programs.

A collection of papers (41) presented at the 1960 Symposium on Psychophysiological Aspects of Space Flight held at the USAF School of Aviation Medicine surveys the varied research in progress. In a recent volume (122), Sells and Berry bring together papers which demonstrate the contributions of aviation medicine and human factors sciences.

Goshen (60 and 61) clearly delineates between the closely linked fields of psychological and physiological research in space medicine. Other views which are important for an understanding of the strictly psychological problem areas are presented by Hauty (64, 65 and 66), Michael (96), Peters (108) and Mundy-Castle (100).

The limited extent of Russian interest in space psychology is indicated by the scarcity of published material which can be found from that source. Although Russian documentation in space technology and the physiology of space medicine is plentiful, relatively scant attention is paid to the psychological problems surrounding man in space. Shternfeld (124) includes one brief chapter on man in space in his *Soviet Space Science*. Platonov (110) deals with the psychological aspects and emphasizes the value of astronaut training.

Confinement, Isolation and Sensory Deprivation

The nature of confinement stress has been defined to include both isolation and sensory deprivation. According to Flaherty et al. (42), sensory deprivation refers to the overall

Received May 8, 1961.

¹ Reference Librarian and Bibliographer.

² Numbers in parentheses indicate Annotated Bibliography.

decrease in sensory stimulation, and isolation to one consequence of this—separation from more complex, meaningful environmental signals.

Actual experiences of submarine crews (76, 83, 84, 101 and 128) have produced valuable information on the reactions of men confined in cramped quarters for a long period of time. This problem related to space flight crews has been considered in several simulation studies at the Air Crew Equipment Laboratory, U. S. Naval Air Material Center (20, 21, 47, 48 and 62).

Ruff (114) points to the emotional strain of isolation on the space traveler and emphasizes the need for quantity and variety of information inputs. Reports on Arctic expeditions (17, 36 and 99) which isolated man in an unusual environment and indicated the importance of careful evaluation of personnel to withstand such stress have provided factual background material for the space psychologist.

An indication of the "aloneness" man experiences far from Earth has been reported in studies of reactions of pilots at high altitudes. One of these reactions, called the "breakoff phenomenon" or "detachment," is discussed by Clark (27), Clark and Greybiel (28) and Giffen (58).

An extensive amount of research has been conducted in the field of sensory deprivation, beginning with the McGill studies in 1951 (14, 69, 70, 73, 74 and 75). Later research includes that of the Massachusetts Mental Health Center (46 and 71), the Research Center for Mental Health, New York University (59, 77, 78 and 79), the Human Resources Research Office, George Washington University (98), the Psychiatric Research Laboratory, Boston City Hospital (152), and the Aero Medical Laboratory, Wright Air Development Division (32 and 118). A symposium on the subject (132) was held at Harvard Medical School in 1958.

Several bibliographic reports and surveys are of interest here. Wheaton (154) explores the background of sensory deprivation studies and surveys the anecdotal and clinical literature. Scott (120) reviews literature on the intellectual effects of perceptual isolation. Weybrew and Parker (153) have compiled a comprehensive bibliography on sensory deprivation, isolation and confinement.

Weightlessness

Man's proficiency in space will depend to a great extent on his ability to adapt to the weightless state, a condition which has been experienced by subjects for seconds only.

Investigations reported by Gerathewohl (50, 53, 54 and 55), Gerathewohl and Ward (57) and Von Beekh (150) have contributed to the understanding of human reactions at zero-g. Schook (119) and Lansberg (89) note the difficulties of adapting to weightlessness and the training implications.

Psychological Assessment and Training

The psychological assessment and selection of astronauts involves the determination of personality types most suited to the mission and an investigation of the motivation and attitudes of the space candidate.

Slater (130) feels that purely psychological problems do not offer any special difficulties in space flight if judicious choice of personnel is accomplished. Silverman et al. (126) have concluded that dynamic interviews and psychologic testing can weed out those candidates with major defects in ego integration. Gerathewohl (51 and 52), Levine (91) and Ruff and Levy (117) discuss psychiatric and psychological testing methodology used to develop an understanding of the subject's personality.

Candidate evaluation for Project Mercury is summarized by Ruff and Levy (116), Konecki (87) and Wilson (159). The Project Mercury training program is outlined by Newbauer (103). Simons (127) describes testing of candidates for high altitude balloon flights.

Beyer and Sells (15) and Sells and Berry (121 and 123) present a comprehensive analysis of space crew requirements related to psychological adaptability. Flinn (45) appraises the desirable personality characteristics of an effective space man.

Gerathewohl (51 and 52) and Solomon (131) analyze the important role of motivation in the emotional reactions of the space traveler. Nehnevajsa (102), Peters (108) and Rohrer (113) point out that the study of interaction in small groups is important as a basis for selection of space crews of the future.

Because man must be trained for space in a ground environment, his conditioning must be in the form of simulation of situations that he will encounter. Summary descriptions of equipment design (1, 107 and 112) are good basic reading for a consideration of the many stresses which must be simulated.

An evaluation of the role of simulators in training is given by Eckstrand and Rockway (35), Balke (11) and Gerathewohl (56) review several significant experimental studies. Other simulation experiments at the USAF School of Aviation Medicine are reported by Hawkins (68), Steinkamp and Hauty (136), Steinkamp and Hawkins (137), Steinkamp et al. (138) and Strughold (141).

Fatigue, Performance and Work Load

In maintaining his performance, one major problem the astronaut faces is fatigue resulting from boredom, prolonged commitment to a skilled task, monotonous environment and loss of sleep.

According to Cunningham (31), the welfare of the space passenger and the success of his mission may well depend upon keeping him fully occupied with as great a variety of activities as possible. Jackson (82) points to the need for more consideration of the kind and quantity of controlling work that a pilot can do, and Lowry (94) suggests possible duties of a hypothetical crew.

Experimental studies by Gaito et al. (47) indicate that simple routine tasks should be minimal and should be supplemented with more complex tasks requiring higher level behavior. Payne (106) stresses the importance of work and living space requirements, and Hauty (65) charts operator proficiency and hours of work during simulated flights.

The role of fatigue in relation to work proficiency is dealt with in detailed reports by Hauty and Payne (67), Adams (2) and Kraft (88).

Fitts (40) reviews procedures expected to lessen the effects of adverse environments on skilled performance. Hartman (63) and Miller (97) study the factor of information overload.

Bibliographies and Other Information Sources

Many comprehensive bibliographic reports and current information sources in the broad area of space medicine provide useful reference material for the study of man-in-space psychology.

Aerospace Medicine, monthly publication of the Aerospace Medical Association, includes a regular section, "Abstracts of Current Literature." The National Library of Medicine compiles the monthly *Index Medicus*, which has a subject breakdown "Space Flight—Psychology." The Institute of the Aerospace Sciences, in cooperation with the National Science Foundation and the Air Force Office of Scientific Research, began in January 1961 to issue *International Aerospace Abstracts*, a monthly publication replacing the section "Aerospace Reviews" in *Aerospace Engineering*.

Other valuable sources of information are the *Technical Abstract Bulletin* of the Armed Services Technical Information Agency (ASTIA); *Psychological Abstracts*, the American Psychological Association; *Air University Periodical Index*, Air University Library, Maxwell Air Force Base; and the *Bulletin of the Public Affairs Information Service*. ASTIA

also publishes regular bibliographies of reports received in bioastronautics (5 through 8).

Jet Propulsion Laboratory, Pasadena, Calif., is currently preparing and distributing two monthly publications: *Astronautics Information: Open Literature Survey* and *Astronautics Information: Abstracts*.

The most comprehensive recent bibliographies in the

specific field of space psychophysiology are those by Burns et al. (22), Banghart and Pattishall (12) and Chasen et al. (26). Extensive references to related material may also be found in general space bibliographies (13 and 37), space medicine bibliographies (145 and 148), and an indexed compilation of documentation on the general subject of psychological stress (3 and 4).

Annotated Bibliography

1 Ackerman, M., "An integrated space-flight simulator," Institute of Radio Engineers, *Transactions on Military Electronics*, vol. MIL-3, no. 3, July 1959, pp. 92-98.

Describes a simulator for space flight training designed to integrate the physiological and psychological stresses on the future space crew and provide a complete environment for experimentation. Early phasing of the integrated simulator with the space vehicle is suggested as a better foundation for design of the space cabin or capsule than sole dependence on feedback from early flights.

2 Adams, O. S., "Aircrew fatigue problems during extended endurance flight; Phase I: Planning," Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, WADC-TR-57-610, May 1958, 86 pp.

Discussion of an experimental program to determine the effects of confinement-induced stresses on a five-man crew isolated in a flight station for a period of 120 hr; 129 references.

3 Applezweig, Mortimer H., comp. and ed., "Psychological stress and related concepts: A bibliography," Connecticut College, New London, Conn., *Tech. Rep.* no. 7, AD 158 085, Dec. 1957, 185 pp.

References cover a period of some 30 years and deal with studies of anxiety, ego-involvement, frustration, breakdown, conflict, tension, fatigue, excessive stimulation, understimulation, extremes of deprivation or of environmental conditions, pressure, and emotional conditioning.

4 Applezweig, Mortimer H., comp. and ed., "Psychological stress and related concepts: Indices to bibliography," Connecticut College, New London, Conn., *Supplement to Tech. Rep.* no. 7, AD 158 086, Jan. 1958, 35 pp.

5 Armed Services Technical Information Agency, "Bio-astronautics: An ASTIA report bibliography," Arlington, Va., AD 306 007, Feb. 1959, 9 pp.; secret. Lists 48 references from 1952 to 1958.

6 Armed Services Technical Information Agency, "Bio-astronautics: An ASTIA report bibliography," Arlington, Va., AD 211 775, Feb. 1959, 157 pp.

Lists 919 references. Covers the subject matter through 1958 insofar as report literature, represented by ASTIA holdings, is concerned.

7 Armed Services Technical Information Agency, "Bio-astronautics: An

ASTIA report bibliography," Arlington, Va., Supplement to AD 211 775, AD 233 000, Feb. 1960, 43 pp.

Previous bibliographies have covered the literature on bioastronautics through 1958. This supplemental bibliography brings the subject matter up to date through 1959 insofar as report literature, represented by ASTIA holdings, is concerned.

8 Armed Services Technical Information Agency, "Bio-astronautics: An ASTIA report bibliography," Arlington, Va., AD 315 200, Feb. 1960, 11 pp.; secret.

9 Armstrong, Harry, Haber, Heinz and Strughold, Hubertus, "Aeromedical problems of space travel," *J. Aviation Medicine*, vol. 20, Dec. 1949, pp. 383-417.

Points out that even if the mechanical problems concerned with space flight were solved, there would still remain the physiological-psychological problems of man in space.

10 Askren, W. B., "Man functions in space flight," Wright Air Development Division, Wright-Patterson Air Force Base, Ohio, AD 238 480, 1959, 10 pp.

Presented to the Panel on Psychology of the Armed Forces-NRC Committee on Bio-Astronautics, in Washington, D. C., Dec. 2, 1959. The hypothetical activities of a three-man crew of an Earth-to-moon flight are analyzed and the following selection criteria suggested: sensory abilities, such as vision and hearing; mental abilities, such as general intelligence, computational skills and problem solving ability; motor skills, such as tracking skill and response times; tolerance to physiological stress, such as high g, heat and cold; and adaptation to new experiences (psychological stress), such as confinement, new food forms and crew interaction.

11 Balke, Bruno, "Experimental studies on the conditioning of man for space flight," *Air University Quarterly Review*, vol. 11, no. 1, Spring 1959, pp. 61-74.

Discussion of experimental work on man's conditioning and tolerance limits to some of the anticipated singular stresses, with emphasis upon cross-adaptability for complex stress situations.

12 Banghart, Frank W. and Pattishall, Evan G., "Human factors at extreme altitudes: Synopsis and bibliography," Division of Educational Research, University of Virginia, Charlottesville, Va., AD 242 348, (HQ ARDC-TR-60-7), March 1960, 111 pp.

Lists 1076 references to 1959 in space medicine, ecology, behavior and perform-

ance, acceleration and deceleration, weightlessness, radiation effects, instrumentation, monitoring and communication, selection and training.

13 Benton, Mildred, comp., "The literature of space science and exploration," Naval Research Laboratory, Washington, D. C., *NRL Bibliography* no. 13, AD 210 057, Sept. 1958, 264 pp.

An annotated bibliography of 2274 references from 1903 to June 1958 on all aspects of space flight, rockets and satellites; subject indexed.

14 Bexton, W. H., Heron, W. and Scott, T. H., "Effects of decreased variation in the sensory environment," *Canadian J. Psychology*, vol. 8, no. 2, 1954, pp. 70-76.

Reports on study by the McGill Group; indicates that progressive intellectual deterioration and vivid visual hallucinations result from exposure to a monotonous environment for periods ranging from several hours to six days.

15 Beyer, David and Sells, S. B., "Selection and training of personnel for space flight," *J. Aviation Medicine*, vol. 28, Feb. 1957, pp. 1-6.

An outline of space crew requirements presented as a basis for guiding research in the selection and training of personnel. Includes bibliography.

16 Billingham, J., "Space medicine," *Nature*, London, vol. 182, no. 4650, Dec. 13, 1958, pp. 1645-1646.

Reviews the first symposium on space medicine in Great Britain, held in London, Oct. 16 and 17, 1958. Among papers presented were: "Consequences of Weightlessness" by M. P. Lansberg (89), "Impairment of Human Performance in Control" by K. F. Jackson (82), and "Effects of Sensory Impoverishment, Confinement, and Deprivation of Sleep" by C. Cunningham (31).

17 Boag, T. J., "The white man in the arctic: A preliminary study of the problems of adjustment," *Amer. J. Psychiatry*, vol. 109, 1952, pp. 444-449.

Observations are reported on some of the psychological stresses of life in the Arctic and human reactions to them.

18 Bradney, Pamela J., "Sociological aspects of space travel," *Brit. Interplanet. Soc. J.*, vol. 16, July-Sept. 1957, pp. 173-176.

Reply to Wilcox (*JBIS*, vol. 16, 1957, p. 7) pointing to factors making for group integration and easy interaction on board space ship.

19 Brody, Eugene B., "Discussion," *Amer. J. Psychiatry*, vol. 115, no. 12, June 1959, pp. 1112-1113.

Summary discussion of papers presented at Symposium on Psychiatric Aspects of Space Travel (143), pointing out that the central psychological problem of man's initial ventures into space may be that of separation.

20 Burns, Neal, "Environmental requirements of sealed cabins for space and orbital flights. A second study. Part I. Rationale and habitability aspects of confinement study," Air Crew Equipment Laboratory, U. S. Naval Air Material Center, Philadelphia, Pa., Rep. no. NAMC-ACEL-413, 1 vol., AD 229 908, Dec. 8, 1959.

An introduction is given to a six-part report on a confinement experiment in which six men were isolated for a period of eight days in a chamber of limited space and facilities equipped with a new type of rebreathing system. A description of the test chamber and of the performance measures that were used is given.

21 Burns, Neal and Ziegler, Ralph B., "Environmental requirements of sealed cabins for space and orbital flights—A second study. Part 3. Effects of long term confinement on personality and perception," U. S. Air Crew Equipment Laboratory, U. S. Naval Air Material Center, Philadelphia, Pa., Rep. no. NAMC-ACEL-415, 1 vol., AD 242 158, July 22, 1960.

Six naval enlisted men served as subjects for eight days of confinement in a simulated space vehicle. The effects on personal perception, personality changes and group interaction are described. The data obtained are compared with the pre-confinement observations in order to provide a detailed analysis of the personality and perceptual changes that can be anticipated in future long range space and orbital flights. An extensive review of the anecdotal and experimental literature is also provided.

22 Burns, Neal M., Ziegler, Ralph B., Noble, Rosalie and Gifford, Edmund C., "Environmental requirements of sealed cabins for space and orbital flight: A bibliography of psychophysiological studies relevant to space and orbital flight," U. S. Air Crew Equipment Laboratory, U. S. Naval Air Material Center, Philadelphia, Pa., Rep. no. NAMC-ACEL-441, Oct. 26, 1960.

Lists 582 psychological, physiological, and environmental reports. Survey was completed in April 1960.

23 Burrows, A., "Psychological stress in space travel," *New Scientist*, vol. 3, no. 75, April 1958, pp. 12-14.

24 Cameron, D. Ewen, Levy, Leonard, Ban, Thomas and Rubinstein, Leonard, "Sensory deprivation: Effects upon the functioning human in space systems," *Psychophysiological Aspects of Space Travel* ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 225-237.

Summarizes sensory deprivation studies since early 1950's and presents conclusions.

25 Campbell, P. A. and Gerathewohl, S. J., "The present status of the problems of weightlessness," *Texas State J. Medicine*, vol. 55, no. 4, April 1959, pp. 267-274.

Reports weightlessness orientation studies made by immersing men in water. Man's ability to orient himself depends upon a variety of factors, and during weightless situations the eye becomes the only reliable organ.

26 Chasen, L., Colabrese, E. and Himmelstein, D., "Sealed atmospheres and psychophysiological factors: a bibliography," Missile and Space Vehicle Dept., General Electric Co., Philadelphia, Pa., T.I.S. Rep. R60SD344, AD 238 478, April 1, 1960, 34 pp.

A survey directed toward psychological, physical and biological hazards of space flight; includes literature on the psychological effects of isolation for a prolonged period. Lists 50 references in closed literature; 116 references in open literature.

27 Clark, Brant, "The break-off phenomenon: A feeling of separation from the earth experienced by pilots at high altitudes," United States Naval School of Aviation Medicine, Pensacola, Fla., Research Rep. NM 001 110 100, no. 43, 1956, 6 pp.

Pilots of jet aircraft when flying alone at high altitudes have reported an unusual experience which has been termed "break-off" or physical separation from Earth. This report investigated the occurrence of breakoff phenomenon in 137 jet pilots.

28 Clark, Brant and Greybiel, Ashton, "The break-off phenomenon: A feeling of separation from the earth experienced by pilots at high altitude," *J. Aviation Medicine*, vol. 28, 1957, pp. 121-126.

Analysis of interviews with 137 jet pilots reveals that the breakoff effect, "a condition of spatial orientation in which the pilot conceived himself to be isolated, detached and physically separated from the earth so that he is no longer in contact with it," is a clearly defined phenomenon.

29 Clemedson, Carl-Johan, "Some biophysical and medical problems involved in manned space flight: A review," *Astronautik*, Swedish Interplanet. Soc., no. 1, 1958, pp. 9-36.

Discussion which includes the effects of psychological stress.

30 Cohen, Bertram D., Rosenbaum, Gerald, Dobie, Shirley I. and Gottlieb, Jacques S., "Sensory isolation: Hallucinogenic effects of a brief procedure," *J. Nervous and Mental Disease*, vol. 129, no. 5, Nov. 1959, pp. 486-491.

The study was designed to develop a brief isolation procedure for more convenient exploration of susceptibility to hallucinatory behavior and the variables contributing to this process. The duration of isolation was limited to one hour.

31 Cunningham, Cyril, "The effects of sensory impoverishment, confinement and sleep deprivation," *Brit. Interplanet. Soc. J.*, vol. 17, May-June 1960, pp. 311-313. (Also published as "First man in space will be like a prisoner in solitary confinement," *Space*, Jan.-March 1959, pp. 17-21.)

Paper presented at Space Medicine Symposium, London, Oct. 16-17, 1958. Discusses the probable effects of conditions similar to those expected in space

flight—isolation, restriction of personal movement, deprivation of auditory and external visual stimuli and sleep, and unawareness of the vehicle's movement in space. Conclusions are drawn from experimental tests and from experiences of persons who have spent many months in foreign isolation prisons.

32 Dempsey, C. A., Van-Wart, F. D., Duddy, J. H. and Hockenberry, J. A., "Long term human confinement in space equivalent vehicles," *J. Astronautics*, vol. 4, Autumn 1957, pp. 52-53.

Brief report on Aero Medical Laboratory's 1954 study of 56 hr confinements in fighter cockpits, and description of current 120-hr confinements in space equivalent compartments.

33 Dostal, R. and Kersey, T., "Space medicine," *The Iowa Engineer*, vol. 60, no. 5, Feb. 1960, pp. 18-20.

Explains the psychological effects of confinement, hypoxia, excessive oxygen, ecological food systems, and space hibernation.

34 Duff, F. L. and Collins, T. A., "Spacecrew epidemiology," *ASTRONAUTICS*, vol. 5, no. 2, Feb. 1960, pp. 42, 44-48.

Careful crew selection and isolation training are vital to assure the success of space missions.

35 Eckstrand, G. A. and Rockway, M. R., "The role of simulators for space crew training," *ASTRONAUTICS*, vol. 5, no. 2, Feb. 1960, pp. 38-39, 76-80.

Training simulators designed to assist in the development, maintenance and evaluation of human skills within a space holding facility are discussed. Inter-crew training (system checkout and final crew selection, mission briefing, and establishment of standard operating procedures) is evaluated.

36 Eilbert, L. R. and Glaser, R., "Differences between well and poorly adjusted groups in an isolated environment," *J. Applied Psychology*, vol. 43, Aug. 1959, pp. 271-274.

Explores the possibility of identifying variables which will predict an individual's adjustment to isolated restrictive environment. Describes survey and test instruments used in the development of selection techniques for men on Arctic bases. Discusses relationship of data to other types of isolated environments.

37 Estep, Raymond, "A space bibliography through 1958," Documentary Research Division, Research Studies Institute, Air University, Maxwell Air Force Base, Ala., AU-283-58-RSI, 1959, 109 pp.

Lists 1832 references; includes subject and author index.

38 Evrard, E., Henrotte, J. G. and Jonckheere, P., "Contribution à l'étude du comportement psycho-physiologique d'un sujet isolé en milieu confiné dans des conditions très défavorables," (Contribution to the study of psycho-physiological behavior of an isolated subject confined in a small closed cabin under very unfavorable conditions), *Medicine Aeronautique*, Paris, vol. 14, no. 1, 1959, pp. 31-50. (In French, with English summary.)

Describes test in which 40-year-old male volunteer was isolated for 31 hr in a decompression chamber which was very limited and cramped in space, and which contained only the minimum features of comfort. Periods of depression occurred during the evenings.

39 Faust, H., "Man—The unknown factor in interplanetary flights," *Universitas* (Stuttgart), vol. 2, no. 3, 1958, pp. 289–296. (In English.)

Although external and physiologic dangers can be overcome by providing a suitable cabin for the human occupant, the consequences of the psychologic strain of being confined in an insecure, deprived and limited environment for an extended period of time cannot yet be estimated. The breakdown of one of the crew members may endanger the entire project.

40 Fitts, Paul M., "Skill maintenance under adverse conditions," *Psychophysiological Factors of Space Travel*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 309–322.

Lists eight factors having adverse effect on skilled performance and presents procedures to lessen effects.

41 Flaherty, Bernard E., ed., *Psychophysiological aspects of space flight*, Columbia University Press, N. Y., 1961, 368 pp.

Papers presented by experts of the U. S. space research program at a symposium sponsored by the School of Aviation Medicine, USAF Aerospace Medical Center, May 26 and 27, 1960.

42 Flaherty, Bernard E., Flinn, Don E., Hauty, George T. and Steinkamp, George R., "Psychiatry and space flight," School of Aviation Medicine, USAF Aerospace Medical Center (ATC), Brooks Air Force Base, Texas, *TR 80-80*, AD 245 416, Sept. 1960, 9 pp.

Discusses some of the emotional hazards that can be anticipated in space flight and their possible effects. Reviews results of experimentation in space flight simulation for training.

43 Flickinger, D. D., "Air Force human factors program for developing manned space operations," *Air University Quarterly Review*, vol. 10, no. 2, Summer 1958, pp. 17–28.

Outlines objectives, assignment of research and development tasks, proposed and on-going projects in USAF life sciences (human factors) research concerning man in space.

44 Flickinger, Don, "Biomedical aspects of space flight," *Man in Space: The USAF Program for Developing the Spacecraft Crew*, ed. by Kenneth F. Gantz, Duell, Sloan and Pearce, N. Y., 1959, pp. 41–63.

In addition to biological hazards, enumerates the psychological hazards of space flight. Reviews work in progress in the Air Force Air Research and Development Command laboratories.

45 Flinn, Don E., "Psychiatric factors in astronaut selection," *Psychophysiological Aspects of Space Flight*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 87–95.

A general discussion of astronaut evaluations from the psychological viewpoint.

46 Freedman, S. J. and Greenblatt, Milton, "Studies in human isolation," Massachusetts Mental Health Center, Boston, Massachusetts, *Report for 15 April 1958 to 14 April 1959 on Biophysics Research* (WADC TR-59-266), AD 231 290, Sept. 1959, 46 pp.

Thirty normal college-age subjects were run in eight-hr experimental sessions involving: nonpatterned visual and auditory input combined with social isolation; visual deprivation (blackout) combined with auditory non-patterning and social isolation; and social isolation alone.

47 Gaito, John, Hanna, T. D., et al., "Environmental requirements of sealed cabins for space and orbital flights. Part 3. Performance and habitability aspects of extended confinement," Air Crew Equipment Laboratory, U. S. Naval Air Material Center, Philadelphia, Pa., *Rep. no. NAMC ACEL-385*, AD 205 381, Sept. 26, 1958, 52 pp.

Six men were isolated and confined in a limited space for seven days breathing a relatively high oxygen concentration. Performance on simple, routine psychomotor tasks was variable, whereas that on more complex tasks was consistent and sometimes increased in effectiveness toward the end of the confinement period.

48 Gell, C. F., *Psychophysiological aspects of a multiple crew compartment study*, AMERICAN ROCKET SOCIETY, N. Y., *ARS preprint 681-58*, 1958.

The conclusion of this study was that, given men with good training, discipline and high motivation, confinement similar to that expected in space flight of at least seven days can be well tolerated on a psychophysiological basis. The extensive accumulation of data from this seven-day exposure has been evaluated and compiled in reports by the Air Crew Equipment Laboratory.

49 Generales, Constantine D. J., Jr., "Space medicine and the physician," *N. Y. State J. Medicine*, vol. 60, no. 11, June 1, 1960, pp. 1741–1761.

Reviews the background of man's desire to travel through interplanetary space from year 1500 A.D. Discusses various aspects of space medicine, including psychological problems of weightlessness, isolation and day-night cycle. Lists human factors in space travel which need further research. The task of space medicine is to adjust man to space environmental conditions which affect him physically and psychologically.

50 Gerathewohl, Siegfried J., "Personal experiences during short periods of weightlessness reported by sixteen subjects," *Astronautica Acta*, vol. 2, 1956, pp. 203–217.

Reactions to weightlessness by passengers in a Lockheed T-33 flying a Keplerian trajectory as revealed in Air Force School of Aviation Medicine tests of 1955–1956.

51 Gerathewohl, Siegfried J., "Psychological problems of selection, hold-

ing and care of astronauts," *ASTRONAUTICS* vol. 5, no. 2, Feb. 1960, pp. 36–37, 52–54.

In the recruiting and selection of astronaut, the problem of motivation is of utmost importance. Psychiatric and psychological testing to develop an understanding of the subject's personality is discussed.

52 Gerathewohl, Siegfried J., "Psychological problems of selection, holding, and care of space fliers," U. S. Army Medical Services, Research and Development Command, Bioastronautics Research Unit, Washington, D. C., *Special Rept. CSCRD-16-4*, AD 232 121, Nov. 13, 1959, 13 pp.

The psychological requirements for the selection of space fliers are examined as to their validity. Examples of related activities are presented. The main purpose of preparations must be to establish a natural pattern of conditioning and familiarization, to develop skills and abilities parallel to the development of hardware, and to maintain the interest and motivation of the astronaut.

53 Gerathewohl, Siegfried J., "Weightlessness," *ASTRONAUTICS*, vol. 2, no. 4, Nov. 1957, pp. 32–34, 74–75.

Discussion of the phenomenon and reports on experimental studies of weightlessness conducted at USAF School of Aviation Medicine.

54 Gerathewohl, Siegfried J., "Weightlessness," *Man in Space: The United States Air Force Program for Developing the Spacecraft Crew*, ed. by Kenneth F. Gantz, Duell, Sloan and Pearce, N. Y., 1959, pp. 108–132.

Discusses physical concepts, operational aspects and psychological aspects of weightlessness. Reviews experiments with humans.

55 Gerathewohl, Siegfried J., "Weightlessness: The problem and the Air Force Research Program," *Air University Quarterly Review*, vol. 10, Summer 1958, pp. 121–141.

Includes tables on responses of 47 human subjects to short periods of virtual weightlessness and discussion of psychological aspects.

56 Gerathewohl, Siegfried J., "Work proficiency in the space cabin simulator," *Aerospace Medicine*, vol. 30, Oct. 1959, pp. 722–735.

Reports on two studies made at the U. S. Air Force School of Aviation Medicine: (1) Airman First Class D. G. Farrell and his seven-day stay in a space flight simulator measuring about 96 cubic feet. Kraepelin's work performance test was used in a measure of performance, proficiency, personality variables, and motivation; (2) Dr. Bruno Balke and Master Sergeant S. M. Karst who were sealed in a cabin on Sept. 12, 1958.

57 Gerathewohl, Siegfried J. and Ward, Julian E., "Psychophysiological and medical studies of weightlessness," *Physics and Medicine of the Atmosphere and Space—The Proceedings of the Second International Symposium on the Physics and Medicine of the Atmosphere and Space*, held at San Antonio, Texas, Nov. 10, 11 and 12, 1959, sponsored by the School of

Aviation Medicine, Aerospace Medical Center (ATC), Brooks Air Force Base, Texas, ed. by Otis O. Benson Jr. and Hubertus Strughold, John Wiley and Sons, N. Y., 1960, pp. 422-434.

Reports study of group of 46 men and one woman for weightlessness tolerance at the USAF School of Aviation Medicine. Tabular account of results with psychological reactions.

58 Giffen, Martin B., "Break-off: A phase of spatial disorientation," *U. S. Armed Forces Medical J.*, vol. 10, no. 11, Nov. 1959, pp. 1299-1303.

Reports study made of symptoms described by pilots experiencing the break-off sensation at high altitudes. Phenomena appears to be chiefly one of isolation and can be artificially produced by the removal of sources of stimulus.

59 Goldberger, L. and Holt, R. R., "Experimental interference with reality contact (perceptual isolation): Method and group results," *J. Nervous and Mental Disease*, vol. 127, 1958, pp. 99-112.

60 Goshen, Charles E., "Man in space and psychological research," *ASTRONAUTICS*, vol. 6, no. 3, March 1961, pp. 38-39, 78, 80.

Emphasizes need for true psychological research as opposed to physiological research in the space effort.

61 Goshen, C. E., "Space science and research in human behavior," *Diseases of the Nervous System*, vol. 21, no. 2, supplement, Feb. 1960, pp. 106-111.

62 Hanna, T. D. and Gaito, J., "Performance and habitability aspects of extended confinement in sealed cabins," *Aerospace Medicine*, vol. 31, no. 5, May 1960, pp. 399-406.

Study was undertaken to determine the effects on performance of six men during seven days of confinement in a small area at a simulated altitude of 10,000 ft with an oxygen concentration equivalent to 55% at sea level.

63 Hartman, Bryce O., "Time and load factors in astronaut proficiency," *Psychophysiological Factors of Space Travel*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 278-308.

Considers the astronaut's job, effects of load, long-term effects, and weightlessness. 47 references.

64 Hauty, George T., "Human performance in space," *Man in Space: The United States Air Force Program for Developing the Spacecraft Crew*, ed. by Kenneth F. Gantz, Duell, Sloan and Pearce, N. Y., 1959, pp. 84-108.

Discusses conditions peculiar to a closed ecological system in space, with emphasis on the psychological effects on man of confinement, detachment, sensory deprivation and fatigue. Reviews briefly some of the Air Force studies concerning these factors.

65 Hauty, George T., "Psychological problems of space flight," *Physics and Medicine of the Atmosphere and Space—The Proceedings of the Second International Symposium on the Physics and*

Medicine of the Atmosphere and Space, held at San Antonio, Texas, Nov. 10, 11, and 12, 1959, sponsored by the School of Aviation Medicine, Aerospace Medical Center (ATC), Brooks Air Force Base, Texas, ed. by Otis O. Benson Jr. and Hubertus Strughold, John Wiley and Sons, N. Y., 1960, pp. 409-421.

Discusses in light of initial research the psychological problems of sensory deprivation, detachment, confinement and fatigue. Charts operator proficiency and hours of work during simulated space flights.

66 Hauty, G. T., "Human performance in the space travel environment," *Air University Quarterly Review*, vol. 10, no. 2, Summer 1958, pp. 89-107.

Reviews conditions of a closed ecological space system, limitations of the human operator, selection and training of space pilots and programming of functions.

67 Hauty, G. T. and Payne, R. B., "Fatigue, confinement, and proficiency decrement," *Vistas in Astronautics: Proceedings of the First Annual AFOSR Astronautics Symposium*, ed. by M. Alperin, M. Stern and H. Wooster, Pergamon Press, N. Y., 1958, pp. 304-309.

Reports a study in which subjects were required to perform a monitoring task for 30 consecutive hours. Nearly all subjects experienced perceptual disturbances with adverse effect upon proficiency.

68 Hawkins, Willard R., "Space cabin requirements as seen by subjects in the space cabin simulator," *AMERICAN ROCKET SOCIETY*, N. Y., *ARS preprint 702-58*, Nov. 17, 1958.

Reports the subjective responses and recommendations of subjects who have actually undergone simulated flights in the space cabin simulator at the School of Aviation Medicine during the previous two years. The reactions of various subjects to the conditions imposed on them during flights are presented together with suggestions for future improvements.

69 Hebb, D. O. and Heron, Woodburn, "Effects of radical isolation upon intellectual function and the manipulation of attitudes," *Defence Research Board, Canada, Terminal Report on Conditions of Attitude Change in Individuals*, Rep. no. HR63, Ottawa, Canada, AD 82 022, Oct. 1955, 23 pp.; secret.

70 Hebb, D. O., Heron, W. and Bexton, W. H., "The effect of isolation upon attitude, motivation and thought," *Defence Research Board, Canada, Fourth Symposium, Military Medicine 1, in cooperation with McGill University*, Ottawa, Canada, AD 10 258, Dec. 8-10, 1952, 3 pp.; secret.

71 Held, Richard and White, Burton, "Sensory deprivation and visual speed: An analysis," *Science*, vol. 130, no. 3379, Oct. 2, 1959, pp. 860-861.

Authors (Department of Psychology, Brandeis University, Waltham, Mass.) performed experiments in collaboration with S. J. Freedman, director of isolation studies, Massachusetts Mental Health Center. Speeds of moving objects were markedly underestimated by human ob-

servers after prolonged patternless visual stimulation. Even greater underestimation followed exposure to a "noisy" field; on the other hand, exposure to a hyperstable field caused overestimation. The effects of noise simulate those of deprivation.

72 Henry, James P., Eckstrand, A., et al., "Human factors research and development program for a manned satellite," *Human Factors Directorate, Air Research and Development Command, Washington, D. C., ARDC-TR-57-160*, AD 136 410, Oct. 1957, 4 pp.

Among problems needing further study were isolation and confinement, presentation of information, crew skills, selection, training and motivation.

73 Heron, Woodburn, "The pathology of boredom," *Scientific American*, vol. 196, no. 1, January 1957, pp. 52-56.

Reports experiments on human behavioral effects following prolonged exposure to a monotonous environment. Thinking was impaired, childish emotional responses appeared, visual perception was disturbed and brain wave patterns altered. The study began in 1951, under a grant from the Defence Research Board of Canada with D. O. Hebb (McGill University), B. K. Doane, T. H. Scott, W. H. Bexton and author as collaborators.

74 Heron, Woodburn, Doane, B. K., and Scott, T. H., "Visual disturbance after prolonged perceptual isolation," *Canadian J. Psychology*, vol. 10, no. 1, 1956, pp. 13-18.

Three observers were kept in a monotonous sensory environment for six days. On returning to a normal environment, they experienced the following perceptual disturbances: fluctuation, drifting and swirling of objects and surfaces in the visual field; loss of "position constancy"; distortions of shape; accentuation of after-images; and effects on perception of color.

75 Heron, Woodburn, Bexton, W. H. and Hebb, D. O., "Cognitive effects of a decreased variation to the sensory environment," *The American Psychologist*, vol. 8, 1953, p. 366.

76 Hickey, A. E., Jr., "The nuclear submarine as a manned satellite laboratory," *American Society of Mechanical Engineers, Los Angeles, Calif., Paper no. 59-AV-9*, 1959, 4 pp.

77 Holt, Robert R. and Goldberger, Leo, "Assessment of individual reaction to sensory alteration," *Psychophysiological Factors of Space Flight*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 248-262.

Reports two experiments, the methodology of which offers promise for research in space crew selection.

78 Holt, Robert R. and Goldberger, Leo, "Personological correlates of reactions to perceptual isolation," *Research Center for Mental Health, New York University, N. Y., Report on Biophysics of Flight*, (WADC TR-59-735), AD 234 003, Nov. 1959, 46 pp.

The aspects of personality (independently measured by objective tests and by clinical Q-sort ratings) that were found to

correlate with reactions to eight hours of experimental reality-deprivation (perceptual isolation and sensory deprivation) are described. Fourteen male college students were subjected to an experimental deprivation situation similar to that used in the McGill studies, and from their behavior and verbalization, 14 measures of reaction were derived. It is hoped that these findings may be relevant to the problems of space crew selection.

79 Holt, Robert R. and Goldberger, Leo, "Research on the effects of isolation on cognitive functioning," Research Center for Mental Health, New York University, N. Y., (WADD TR-60-260), AD 240 835, March 1960, 22 pp.

The study was undertaken to compare the effects of 100 gamma of LSD-25 with those of eight hours of perceptual isolation, carried out in the McGill style. Fifteen civilian subjects showed impairment on only one of nine cognitive tests after the isolation period. The conclusion was drawn that states produced by LSD and by eight hours of social isolation plus sensory restriction are subjectively quite different.

80 Imus, Henry A., "Human factors in probing outer space," *Advances in the Astronautical Sciences*, vol. 5 (Proceedings of the Second Western National Meeting of the American Astronautical Society, Los Angeles, Calif., Aug. 4-5, 1959), American Astronautical Society, Inc., Plenum Press, N. Y., 1960, pp. 15-24.

Reviews psychological factors affecting the astronaut and relates these to engineering effort; 50 references.

81 International Astronautical Federation, *Space Flight Problems: Being a Complete Collection of All Lectures Held at the 4th Astronautical Congress, Zurich, 1953*, Swiss Astronautical Society for the International Astronautical Federation, Switzerland, Laubscher and Cie, Biel, Switzerland, 1954, 224 pp.

See articles: S. Gerathwohl, Zur frage der orientierung im schwebefreien zustand (The problem of orientation in the gravity-free state). H. J. A. von Beckh, Untersuchung über schwerelosigkeit an versuchspersonen und tieren während des lotrechten sturzfluges (Investigations of effects of weightlessness on research personnel and animals during vertical flight).

82 Jackson, K. F., "Impairment of human performance in control," *Brit. Interplanet. Soc. J.*, vol. 17, May-June 1960, pp. 301-303.

Two experiments in determining the controlling work that a pilot can do and the conditions which affect the quantity and quality of a pilot's work are described. The relationship to problems in space vehicles—size of crews, optimal lengths of watches and fatigue—is emphasized.

83 Kinsey, Jack L., "Psychiatric studies on operation hideout," Naval Medical Research Laboratory, New London, Conn., *NMRL Rep. no. 230*, AD 64 665, July 20, 1953, 12 pp.; confidential.

84 Kinsey, J. L., "Psychologic aspects of the 'Nautilus' transpolar cruise," *U. S. Armed Forces Medical J.*, vol. 10, 1959, pp. 451-462.

85 Konecni, E. B., "Human factors and space cabins," *ASTRONAUTICS*, vol. 3, Jan. 1958, pp. 42-43.

True design criteria for any component or system must include the human element as part of a system.

86 Konecni, E. B., "Human factors in space flight," *Aero/Space Engng.*, vol. 17, no. 6, June 1958, pp. 34-40, 48.

Discusses briefly some problem factors of man in space. Table outlines the possible human factors problem areas in space flight, including psychological and social.

87 Konecni, E. B., "Manned space cabin systems," *Advances in Space Science*, ed. by Frederick I. Ordway, III, Academic Press, N. Y., 1959, pp. 159-265.

Reviews the fundamental requirements for manned space flight, and the possible physiological, psychological and human engineering problems to be encountered. Under psychological factors emphasizes crew selection, confinement, isolation and performance. Outlines the examinations and psychological tests given to selected Project Mercury astronauts; 138 references.

88 Kraft, J. A., "Measurement of stress and fatigue in flight crews during confinement," *Aerospace Medicine*, vol. 30, June 1959, pp. 424-430.

USAF-supported discussion of problem areas which must be explored and description of elaborate facility for conducting a variety of human factors experiments. Among psychological problems considered are: loneliness; confinement and isolation and their effects on performance and behavior; and interpersonal relationships in small groups under stress.

89 Lansberg, M. P., "Cockpit of first spaceship may be like a padded cell," *Space*, Jan.-March 1959, pp. 25-27.

Reprint of paper presented at British Interplanetary Society Symposium, London, 1958. Discusses problems arising from continual weightlessness. These include: air circulation can no longer be induced by convection; body perception will lack gravitational clues; and all muscular activity intended to move an object will have to be relearned.

90 Lawton, R. W., Ekberg, D. R. and Smith, B. J., "Bioengineering problems in early manned space flight," *Annals of the New York Academy of Sciences*, vol. 84, no. 2, Feb. 1960, pp. 27-74. (Abstracted in *Aerospace Medicine*, vol. 31, no. 6, June 1960, p. 525.)

The possible solutions to engineering problems associated with biologic, psychological and medical criteria in manned orbital flight are reviewed.

91 Levine, A. S., "Psychometric considerations in selecting personnel for unusual environments," *Personnel Psychology*, vol. 13, no. 3, Autumn 1960, pp. 233-243.

A discussion of the methodological considerations which underlie the selection of men for unusual environments such as may be encountered in high altitude and space vehicles, long-submergence submarines, and polar regions.

92 Levy, Edwin Z., Ruff, George E. and Thaler, Victor H., "Studies in human isolation," *J. Amer. Medical Assoc.*, vol. 169, Jan. 17, 1959, pp. 236-239.

The study of human isolation confronts workers in the field with an almost unlimited number of variables and parameters to be controlled and considered, including details of the isolation, the concepts and feelings of the experimenters, and all aspects of the personality of the subjects.

93 Lilly, J. C., "Mental effects of reduction of ordinary loads of physical stimuli on intact, healthy persons," *Psychiatric Research Reports*, vol. 5, June 1956, pp. 1-9.

Subjects were suspended in water with a temperature of 94.5 F. They wore blacked out masks and were immersed in the tank except for tops of their heads.

94 Lowry, R., "Selection of personnel for manned space stations," *Aero/Space Engng.*, vol. 19, no. 5, May 1960, pp. 32-33.

A program for selection and training of crewmen is proposed. Training should familiarize the individual with the space program and the vehicle, and coordinate members as a composite crew.

95 McRuer, D. T., Ashkenas, I. L. and Krendel, E. S., "A positive approach to man's role in space," *Aero/Space Engng.*, vol. 18, no. 8, Aug. 1959, pp. 30-36.

Discussion emphasizing the reliability improvement afforded by man's ability at failure detection, replacement and repair. Among negative factors investigated are respiration, food, environment and social factors. Suggestions are made for research of possible behavior problems.

96 Michael, D. N., "How to keep space crews content," *Missiles and Rockets*, vol. 3, April 1958, pp. 110-114.

The social and psychological circumstances and consequences which may affect space crew operations are discussed. Some of the problems to be resolved are those associated with perception of unfamiliar sensations including weightlessness and a general paucity of stimuli, interaction of personalities of crew members in confinement, stress due to knowledge of probability of disaster, effects of boredom, and crew selection and training.

97 Miller, James G., "Sensory overloading," *Psychophysiological Aspects of Space Travel*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 215-224.

Describes a research program at the Mental Health Research Institute of the University of Michigan dealing with overloading the astronaut with information input.

98 Monty, Richard A., Myers, Thomas and Murphy, Donald B., "Effects of correct and incorrect knowledge of results on ability to count auditory stimuli. Task ENDORSE I," U. S. Army Leadership Human Research Unit, Human Resources Research Office, George Washington University, Washington, D. C., *Research Rep. 3*, March 1960.

The purpose of this study was to develop a measure which would be useful in detect-

ing changes both in utilization of correct information and in susceptibility to misinformation under conditions of partial or complete sensory deprivation.

99 **Mullin, Charles S. and Connery, H. J. M.**, "Psychological study at an Antarctic IGY station," *U. S. Armed Forces Medical J.*, vol. 10, no. 3, March 1959, pp. 290-296.

Thirty-nine men participating in the International Geophysical Year research program for one year at an isolated Antarctic station were studied to determine the effectiveness of the Navy's screening program, and to study group and individual reactions to disappointment and disillusionment, stress of isolation, boredom, insomnia, anxiety and hysteria. Also included were studies on group-leader tension, cohesiveness, subgroup expectations.

100 **Mundy-Castle, C. A.**, "Psychological problems of space flight," *South African J. Sci.*, vol. 54, no. 9, Sept. 1958, pp. 225-230. (Abstracted in *Aerospace Medicine*, vol. 31, no. 3, March 1960, p. 264.)

A discussion primarily of the effects of perceptual isolation to which a space traveler may be subjected in a spacecraft as a result of monotony or inadequate variation in his sensory influx.

101 **National Research Council, Panel on Psychology and Physiology**, "A survey report on human factors in undersea warfare," Committee on Undersea Warfare, National Research Council, Washington, D. C., 1949, 541 pages.

102 **Nehnevajsa, Jiri**, "Sociology and the space age," *ASTRONAUTICS*, vol. 4, Sept. 1959, pp. 36-37.

Points to new areas of research opening up for sociologists in the space age—among them the contributions of the sociology of small groups, the selection of astronauts, and particularly the selection of crews. Stresses importance of major simulation facility for systematic study of problems related to crew selection and behavior during spaceflight.

103 **Newbauer, John**, "Training an astronaut," *ASTRONAUTICS*, vol. 4, no. 9, Sept. 1959, pp. 31, 41.

Chart outlines astronaut activities in preparation for Project Mercury flights.

104 **Ogle, Dan C.** "Man in a space vehicle," *United States Armed Forces Medical J.*, vol. 8, 1957, pp. 1561-1570.

Despite many engineering achievements over the centuries, man himself has remained unchanged. In space he will face problems of excessive accelerative forces, cosmic radiation, and the psychological strains of existing in a monotonous lonesome void.

105 **Page, J. D.**, "Kayak hunting and space flight," *American Psychologist*, vol. 14, no. 10, Oct. 1959, p. 655.

Compares the psychologic stresses of kayak hunting (confinement, sensory restriction) with those of space flight, including the possible hypnotic effects.

106 **Payne, F. A.**, "Work and living space requirements for manned space stations," *Aero/Space Engng.*, vol. 19, no. 5, May 1960, pp. 34-35.

Adequate spatial accommodations for sleeping, eating, cooking, recreation, medical facilities, and clerical activities must be combined with attractiveness of surroundings in order to counteract the monotony and boredom which derive from isolation with an unchanging set of people.

107 **Pecoraro, J. N.**, "The Navy's role in astronautical training devices," *Military Systems Design*, vol. 4, no. 4, July-Aug. 1960, pp. 24-27.

Describes the crew trainer systems and equipment under consideration by the U. S. Naval Training Device Center for simulating orbital and space flight.

108 **Peters, G. A.**, "Psycho social problems of manned spaceflight," *ASTRONAUTICS*, vol. 5, no. 3, March 1960, pp. 30-31, 89-92.

Analysis of work done to date on human factors indicates the need for a full-scale study of the psychological, social and sexual problems of spaceship crew members. An important variable to be considered is social interaction between a small group in very close contact for an extended period of time. Points to questions which should be investigated: What is the desirable mixture of crew members in terms of age, sex, experience, education and socioeconomic background? What is the optimum number of crew members to reduce the feeling of isolation and lack of communication with Earth? In terms of an appropriate combination of crew personality factors, what are the desirable personality structures for various crew sizes, missions and flight durations? Includes flow diagram of man-machine interaction in manned spaceflight.

109 **Peters, George A. and Dendl, Hans J.**, "Personnel requirements in manned space vehicles," *Astronautical Sciences Review*, vol. 2, no. 3, July-Sept. 1960, pp. 9-12.

Describes some of the basic considerations which are involved in manning estimates (optimum size of space crew) and crew scheduling for extended flights, including customary temporal patterns, sleep-wakefulness cycles, sustenance, human work performance, control of scheduling. Complications may result from simulation of space flight conditions for research or training because test subjects may experience chronic or acute psychological impairment.

110 **Platonov, K. K.**, "Psikhologicheskie problemy kosmicheskogo polëta," *Vosprosy Psikhologii* (Moskva), vol. 5, no. 3, 1959, pp. 56-65. (In Russian). German translation: "Psychologische Probleme des Raumfluges." Sowjetwissenschaft: Naturwissenschaftliche Beiträge (Berlin), vol. 1959, no. 12, 1960, pp. 1213-1222. English translation: "Psychological Problems of Outer Space Flight." Library of Congress no. 60-13520, 16 Dec. 1959, 12 pp.

The psychologic effects of the conditions of space flight are reviewed with reference to published reports on the reactions of experimental subjects to these conditions. It is concluded that the adverse effects of acceleration, weightlessness, confinement, isolation, and exposure to danger may satisfactorily be counteracted by the proper

conditioning and motivation of space pilots; 21 references.

111 **Pournelle, J. E.**, "Flight to the moon—human crews for space ships," *ASTRONAUTICS*, vol. 3, April 1958, pp. 32-35, 81.

An aviation psychologist notes that, while the isolation involved in interplanetary flight will produce extraordinary stresses, man has often conquered similar situations in the past.

112 **Rabideau, G. F. and Schloredt, D. L.**, "Space systems training devices," Society of Automotive Engineers, Inc., N. Y., Paper-245D, 1960.

Presented at the SAE National Aeronautic Meeting, Los Angeles, Calif., Oct. 10-14, 1960. Discusses a number of hypotheses concerning the desirable characteristics of space crew training and training devices, and introduces principles to facilitate the design and development of useful training curricula and training equipment for manned space systems.

113 **Rohrer, John H.**, "Interpersonal relationships in isolated small groups," *Psychophysiological Aspects of Space Travel*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 263-271.

An analysis of isolation which defines terms and explores motivation and phases of adjustment.

114 **Ruff, G. E.**, "Isolation," *ASTRONAUTICS*, vol. 4, no. 2, Feb. 1959, pp. 22-23, 110-111.

Isolation as source of stress in space flight is discussed in relation to the following variable aspects: psychologic distance ("aloneness"), isolated group conflicts, reactions to confinement, and behavioral changes which accompany sensory deprivations.

115 **Ruff, George E.**, "Experimental studies of stress in space flight," *Amer. J. Psychiatry*, vol. 115, no. 12, June 1959, pp. 1109-1110.

116 **Ruff, G. E. and Levy, E. Z.**, "Psychiatric evaluation of candidates for space flight," *Amer. J. Psychiatry*, vol. 116, Nov. 1959, pp. 385-391. (Abstracted in *Aerospace Medicine*, vol. 31, no. 6, June 1960, p. 523.)

Procedures for selection and psychologic evaluation of the pilots for Project Mercury are described.

117 **Ruff, George E. and Levy, E. Z.**, "Psychiatric research in space medicine," *Amer. J. Psychiatry*, vol. 115, no. 9, March 1959, pp. 793-797. (Abstracted in *Aerospace Medicine*, vol. 31, no. 6, June 1960, p. 522.)

Studies at Aero Medical Laboratory, Wright-Patterson Air Force Base, on the effects of isolation, confinement and sensory deprivation were conducted to investigate the problems of crew selection and existence in an artificial environment. These studies suggest that meaningful structuring of experience is necessary for adaptation to drastic alterations in environments. The conclusion of the authors is that the best qualification for a space crew member is a sound ego.

118 **Ruff, G. E., Levy, E. Z., and Thaler, V. H.**, "Studies of isolation and

confinement," *Aerospace Medicine*, vol. 30, Aug. 1959, pp. 599-604.

Description of confinement tests conducted under simulated operational conditions with crew members of future aircraft and space vehicles, by planning groups of five men in a 17 X 7 X 6-foot compartment area. An evaluation of results is included.

119 Schock, G. J. D., "Sensory reactions related to weightlessness and their implications to space flight," Air Force Missile Development Center, Holloman Air Force Base, New Mex., AFMDC-TR-58-8, AD 135 012, April 1958, 10 pp.

Discussion of the implications of a sensory-starved environment and comparison to the conditions that will prevail in actual space flight. Recommendations for training for future space flights are presented.

120 Scott, T. H., "Literature review of the intellectual effects of perceptual isolation," Defence Research Board, Canada, Rep. no. HR 66, AD 147 724, July 1957, 13 pp.

Reviews evidence from experimental studies regarding efficiency in relation to variability of the sensory environment, problems in explanation of performance change, and habituation with the repetition of stimuli; 93 references.

121 Sells, S. B. and Berry, C. A., "Human requirements for space travel," *Air University Quarterly Review*, vol. 10, Summer 1958, pp. 108-120.

Selection, indoctrination and training procedures for space crews. Aptitude and skill requirements, biologic, medical and physical requirements, and psychologic adaptability requirements; includes bibliography.

122 Sells, S. B. and Berry, Charles A., eds., *Human Factors in Jet and Space Travel: A Medical Psychological Analysis*, Ronald Press, N. Y., 1961, 443 pp.

Professional source book covering the methods and achievements of aviation medicine. Includes sections on the selection and training of crews and the abnormal physical and mental strains of space travel.

123 Sells, S. B. and Berry, Charles A., "Human requirements for space flight," *Man in Space: The United States Air Force Program for Developing the Spacecraft Crew*, ed. by Kenneth F. Gantz, Duell, Sloan and Pearce, N. Y., 1959, pp. 161-177.

Includes summary of psychologic adaptability and indoctrination and training requirements for man in space.

124 Shternfeld, Ari, "Man in cosmic space," in *Soviet Space Science*, Basic Books, N. Y., 1959, pp. 163-189. (Translated from the Russian by Technical Documents Liaison Office, Wright-Patterson Air Force Base, Ohio.)

Presents an excellent, simplified discussion of work and life under conditions of weightlessness. Defines G-force for the layman.

125 Silverman, A. J., Cohen, S. I. and Shmavonian, B., "Investigations of pos-

sible importance in space psychiatry," *Diseases of the Nervous System*, vol. 21, no. 2, supplement, Feb. 1960, pp. 102-105.

126 Silverman, A. J., Cohen, S. I. and Shmavonian, B. M., "Selection techniques for space crews," *Amer. J. Psychiatry*, vol. 115, no. 12, June 1959, pp. 1110-1112.

Describes suggested battery of tests to screen space crews.

127 Simons, D. G., "Psychophysiological aspects of Manhigh," *ASTRONAUTICS*, vol. 4, no. 2, Feb. 1959, pp. 32-33, 62-63.

Describes testing of candidates for Manhigh III, including psychiatric evaluation and isolation tests.

128 Skobba, Joseph S., "Military psychiatry," *Amer. J. Psychiatry*, vol. 116, no. 7, Jan. 1960, pp. 651-653.

A brief summary is presented of the subject areas treated in military psychiatric research during the year 1959, including studies of isolation and the psychologic aspects of the Transpolar Cruise of the Nautilus.

129 Slater, E. T. O., "Psychological consequences of space travel," *Brit. Interplanet. Soc. J.*, vol. 16, no. 3, July-Sept. 1957, p. 176.

Reply to E. J. Wilcox (*JBIS*, vol. 16, Jan.-March 1957, pp. 7-10.)

130 Slater, E. T. O., "Psychological problems of space-flight," *Brit. Interplanet. Soc. J.*, vol. 9, no. 1, Jan. 1950, pp. 14-18.

An application of psychological findings from wartime studies of men under battle stress to the situation in the cabin of a space ship.

131 Solomon, Philip, "Motivations and reactions in space flight," *Psychophysiological Factors of Space Flight*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 272-277.

Discusses in speculative terms the role of motivation and the more dangerous emotional reactions of the space traveler.

132 Solomon, Philip, Kubzansky, Philip, Leiderman, Herbert, Mendelson, Jack H., Trumbull, Richard and Wexler, Donald, *Sensory deprivation: A Symposium held at Harvard Medical School*, Harvard University Press, 1961, 272 pp.

Authorities in their fields develop in papers research pertinent to clinical medicine, public health industry, the elucidation of the phenomenon of brainwashing and the psychological problems of space flight.

133 Solomon, Philip, Leiderman, P. Herbert, Mendelson, Jack and Wexler, Donald, "Sensory deprivation: A review," *Amer. J. Psychiatry*, vol. 114, Oct. 1957, pp. 357-363.

A critical review of some of the most pertinent autobiographical, "brainwashing," and experimental data in the field of perceptual and sensory deprivation. Although there are many separate factors operating in these various situations, it is clear that the stability of man's mental state is dependent on adequate perceptual contact with the outside world. Points out that future studies in this area may well contribute to our knowledge of the

psychological and behavioral patterns of man under conditions of normality and stress.

134 Steinberg, I. I., "Medical electronics: Black bag in space," *ASTRONAUTICS*, vol. 4, no. 6, June 1959, pp. 26-27.

Description of specially designed electronic devices to determine the physiological and psychological responses of man in space.

135 Steinkamp, G. R., "Human factors in space flight," *J. Arkansas Medical Soc.*, vol. 56, Feb. 1960, pp. 338-342.

A discussion of the human factors to be considered in the planning of manned space flight, with reference to the work of the Department of Space Medicine of the USAF School of Aviation Medicine. Topics include the time element, weightlessness, and the psychologic selection and training of space crews.

136 Steinkamp, George R. and Hauty, George T., "Simulated space flights," *Psychophysiological Aspects of Space Flight*, ed. by Bernard E. Flaherty, Columbia University Press, N. Y., 1961, pp. 75-79.

Summarizes reactions of subjects to 30-hr confinement in space cabin simulator.

137 Steinkamp, George R. and Hawkins, Willard R., "Medical experimentation in a sealed cabin simulator," *Physics and Medicine of the Atmosphere and Space-Proc. Second International Symposium on Physics and Medicine of Atmosphere and Space*, San Antonio, Texas, Nov. 10-12, 1959, sponsored by School of Aviation Medicine, Aerospace Medical Center (ATC), Brooks Air Force Base, Texas, ed. by Otis O. Benson Jr. and Hubertus Strughold, John Wiley and Sons, N. Y., 1960, pp. 370-375.

Briefly describes methods for simulating space flight. Points out that group dynamics will be a significant problem as a result of confinement and detached isolation. Studies indicate that man is adaptable and has endurance and strength enough to withstand anticipated stresses, and that through training in sealed cabins he can attain acceptable performance levels.

138 Steinkamp, George R., Hawkins, Willard R., Hauty, G. T., Burwell, R. R., and Ward, J. E., "Human experimentation in the space cabin simulator. Development of life support systems and results of initial seven-day flights," *Air Force, School of Aviation Medicine, Brooks Air Force Base, Texas, Rep. no. 59-101*, AD 226 790, Aug. 1959, 88 pp.

Human experimentation in a small space cabin simulator has been conducted at the USAF School of Aviation Medicine, and the first series of a group of significant studies is presented. The involvement of the simulator, as currently used, is described. Results of a study of five subjects committed to seven-day periods in this small cabin and on a 4:4 work-rest schedule are shown and discussed from the standpoint of the physical, physiologic, psychophysiology and nutritional implications.

139 Stewart, W. K., "Behavior in conventional and extra-terrestrial flight.

Some future aspects of aviation medicine," *Nature*, London, vol. 184, Aug. 22, 1959, pp. 578-582.

Reviews physiological, psychologic and biochemical aspects of behavior and reactions of man and animals during conventional and space flight.

140 Stewart, W. K., "Man as a controller of flight vehicles," *Proc. Royal Instit. Great Britain*, London, vol. 37, pt. 5, no. 169, 1959, pp. 512-523.

Development of modern high-speed, high-altitude aircraft has brought new physiologic problems to human piloting. Also new emphasis has been placed on the psychologic implications of the pilot-instrument system relationship in which reaction time and the role of fatigue are primary factors.

141 Strughold, H., "The U. S. Air Force experimental sealed cabin," *J. Aviation Medicine*, vol. 27, no. 1, 1956, pp. 50-52.

142 Strughold, H. and Benson, O. O., Jr., "Space medical research," *New England J. Medicine*, vol. 261, no. 10, Sept. 3, 1959, pp. 494-502.

Four broad research areas are discussed: biophysics, or ecology of space; the space cabin environment; biodynamics in space flight; and the problem of selection and training of astronautic candidates.

143 Symposium on space psychiatry, *Amer. J. Psychiatry*, vol. 115, no. 12, June 1959, pp. 1108-1112.

Abstracts of papers presented at Psychiatric Aspects of Space Travel (19, 115, 126 and 158), Southern Divisional Meeting of the American Psychiatric Association, Miami, Fla., Dec. 2, 1958.

144 U.S.A.F., School of Aviation Medicine, "Epitome of space medicine," Randolph Air Force Base, Texas, 1 vol., AD 159 052, 1958.

Collection of 30 published articles and research reports, including those on subjects of weightlessness, psychology and day-night cycle.

145 U. S. National Library of Medicine, "Bibliography of space medicine," compiled by Charles A. Roos, U. S. Public Health Service, Washington, D. C., Publication no. 217, *Bibliography Series* no. 21, 1958, 49 pp. (Also published in *U. S. Armed Forces Medical J.*, vol. 10, Jan.-July 1959, pp. 173-217.)

381 references through 1958.

146 U. S. Senate Committee on Aeronautical and Space Sciences, "Space research in the life sciences: An inventory of related programs, resources, and facilities, July 15, 1960," 86th Congress, 2nd Session, Superintendent of Documents, Washington, D. C., 1960, 269 pp.

Bibliography, tables, National Aeronautics and Space Administration organization, agencies of the U. S. Air Force, Navy and Army involved in space research, current major programs and manpower. Appendixes contain: "From Aviation Medicine to Space Medicine" (Hubertus Strughold), 55 references, pp.

229-238; "Impact of Space Exploration on Biology and Medicine" (Clark T. Randt), pp. 239-242; "Psychiatric Evaluation of Candidates for Space Flight" (George E. Ruff and Edwin Z. Levy), pp. 243-249; "Man Alive in Outer Space" (Thomas R. A. Davis), pp. 250-259; "Man in Space...Where We Stand" (Paul A. Campbell), pp. 260-262.

147 Vaeth, J. G., "Training for space," *ASTRONAUTICS*, vol. 1, Fall 1954, pp. 1-6.

The role of training in the preparation of space flight crew members and the need for a space ship simulator to provide realistic training situations.

148 Van Woerden, J., "Space medicine bibliography (Ruimtevaartgeneeskunde)," *Technisch Documentatie en Informatie Centrum voor de Krijgsmacht*, Netherlands, Rep. no. TDCK-16903, AD 227 817, Feb. 1959, 45 pp.; 110 references.

149 Vernon, Jack A., McGill, Thomas E. and Schiffman, Harold, "Visual hallucinations during perceptual isolation," *Canadian J. Psychology*, vol. 12, no. 1, 1958, pp. 31-34.

A total of 20 subjects were confined for periods of up to 72 hr under two conditions of sensory deprivation, one more rigorous than the other. The greatest number and variety of hallucinations were produced in confinement permitting a great deal of visual stimulation, but not pattern vision.

150 Von Beckh, H. L., "Weightlessness and space flight," *ASTRONAUTICS*, vol. 4, no. 2, Feb. 1959, pp. 26-27, 84-86.

Experiments indicate that astronauts' difficulties will not lie in the weightlessness state itself, but in its aggravation of other conditions.

151 Vosburg, Robert, Fraser, Norman and Guehl, John, Jr., "Imagery sequence in sensory deprivations," *Amer. Medical Assoc. Archives of General Psychiatry*, vol. 2, no. 3, March, 1960, pp. 356-357.

Subjects were isolated for three hr in successive conditions of total silence and darkness, total darkness with an ambiguous noise source, and total silence with homogeneous visual-field illumination. Behavior cycle included a period of adaptation characterized by recollection of memories associated with the available sensory input, followed by the development of classic neurotic behavior, with decreasing associative ability, and finally by panic.

152 Wexler, Donald, Mendelson, Jack, Leiderman, P. Herbert and Solomon, Philip, "Sensory deprivation, a technique for studying psychiatric aspects of stress," *Amer. Medical Assoc. Archives of Neurology and Psychiatry*, vol. 79, 1958, pp. 225-233.

As a means of inducing stress, 17 normal male volunteers were exposed to perceptual and sensory deprivation through the use of a tank-type respirator for periods up to 36 hr. Two cases are described in detail. The findings are discussed in relation to pertinent psychological and psychiatric literature.

153 Weybrew, Benjamin B. and Parker, James W., comp., "Bibliography of sensory deprivation, isolation and confinement," Naval Medical Research Laboratories, New London, Conn., *Memo Rep. no. 60-1*, Jan. 1960, 12 pp. (Also published in *U. S. Armed Forces Medical J.*, vol. 11, AD 234 102, Aug. 1960, pp. 903-911.)

Listed are 146 references under headings of: review articles, anecdotal literature, experimental literature, studies of confinement peculiar to space flight; studies of confinement peculiar to submarine environment; sociological and prison confinement; animal studies; miscellaneous.

154 Wheaton, Jerrold L., "Fact and fancy in sensory deprivation studies," School of Aviation Medicine, Brooks Air Force Base, Texas, *Aeromedical Reviews*, Rev. no. 5-59, AD 226 325, Aug. 1959, 60 pp.

Contents: background of the sensory deprivation research; the nature of isolation; review of anecdotal and clinical literature on isolation; review of experimental literature; summary of present knowledge of isolation effects on human behavior; implications for military problems.

155 White, Stanley C., "Human factors and bio-astronautics," *ASTRONAUTICS*, vol. 4, no. 11, Nov. 1959, pp. 34, 166.

State of the art report for the AMERICAN ROCKET SOCIETY Human Factors and Bio-Astronautics Committee; 40 references.

156 Wilcox, E. Jack, "Psychological consequences of space travel," *Brit. Interplanet. Soc. J.*, vol. 16, Jan.-March 1957, pp. 7-10.

Discusses the emotional pressure of outer space, and the prevention and treatment of emotional problems which will arise among space crews.

157 Williams, K. G., *The New Frontier: Man's Survival in the Sky*, William Heinemann Medical Books, London, 1959, 161 pp.

Of special interest are chapters on the basic man-machine control system; the problem of input, man vs. automatic control; fatigue, prevention of fatigue; selection of the man, teamwork in flying; problems and psychological dangers of weightlessness.

158 Williams, R. L., "The nature of space flight," *Amer. J. Psychiatry*, vol. 115, no. 12, June 1959, pp. 1108-1109.

An abstract of a speech presented at the Symposium on Psychiatric Aspects of Space Travel, Miami, Fla., Dec. 2, 1958.

159. Wilson, Charles L., "Project Mercury candidate evaluation program," Aerospace Medical Laboratory, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, *TR-59-505*, AD 234 749, Dec. 1959, 133 pp.

Chapter VII includes a description of psychological tests administered to Project Mercury candidates. The general characteristics desirable in space pilots are listed. In the final candidate recommendation, psychological attributes outweigh physiological attributes.

Feasibility of Turbulent Vortex Containment in the Gaseous Fission Rocket

M. L. ROSENZWEIG¹

W. S. LEWELLEN¹

J. L. KERREBROCK²

Space Technology Laboratories,
Inc.
Los Angeles, Calif.

A semi-empirical extension of the laminar diffusion analysis of Kerrebrock and Meghreblian is made incorporating the experimental information reported by Keyes, concerning the turbulence levels and attainable vortex strengths in jet driven vortex tubes. It is assumed that the turbulent eddy diffusivity increases the diffusion due to concentration gradients, but has no effect on the diffusion due to pressure gradients. From the experimental data and this assumption, it is deduced that satisfactory containment of plutonium, or one of its compounds, in hydrogen is possible for certain ranges of the vortex tangential and radial Reynolds numbers. The restrictions imposed on the system by the requirements for nuclear criticality and fission fragment absorption are examined. It is concluded that a vortex reactor providing a significant increase in gas enthalpy over that of the solid fuel nuclear rocket is feasible, however the total system mass is very large, and the thrust to weight ratio is small. Several schemes for increasing the vortex strengths, and reducing the turbulence level, are reviewed in the light of these results.

IT HAS been recognized for some time that, if the full potential of nuclear energy for rocket propulsion is to be realized, the limitations currently imposed by material temperatures must be circumvented. The performance of a heat transfer rocket such as envisioned by the "Rover" program is limited by the maximum temperature that can be tolerated by the fuel elements, and not, as in chemical rockets, by the energy content of the fuel.

Nuclear rocket performance can be improved if ways can be found to transfer energy from the fuel to the propellant without the use of heat transfer surfaces, and to then isolate the hot propellant from solid boundaries, so that as little energy as possible is deposited in the solid. The first requirement can be met by allowing the propellant and fissionable material to mix in the gas phase, so that the energy transfer takes place by molecular collisions. However, since fissionable material is very costly, and at least a few kilograms of it must be retained in the reactor to insure criticality, some means must be found for containing the fissionable material within the reactor as the propellant passes through it.

One method of containment was considered by Kerrebrock and Meghreblian (1).³ It utilizes the strong radial pressure gradient of a vortex to induce diffusion between the heavy fissionable gas and the light propellant gas. Their analysis assumed laminar, inviscid flow. It showed that the radial mass flow rate per unit of vortex tube length, which is limited by the diffusion rate, is small but independent of the vortex diameter for a given tangential Mach number. This fact led to the concept of a vortex reactor composed of many vortex

tubes of small diameter, and therefore capable of a high total mass flow and a large thrust to weight ratio.

An experimental study of the feasibility of generating vortexes of the character required for this device has subsequently been carried out by Keyes (2). These experiments show first that the flow in cylindrical vortex tubes driven by tangential jets is turbulent, and second that the turbulent shear at the boundary of the vortex is so large as to make the generation of vortexes with the required combination of low radial mass flow and high tangential velocity very difficult. Experiments by Ragsdale (3) have further confirmed these results.

In the light of these experimental data, it seems worthwhile to re-evaluate the feasibility of vortex containment for the gaseous fission rocket. This is the aim of the present study.

Whether containment can be achieved in simple, jet-driven vortex tubes of the type studied by Keyes is considered in the first part of this paper. Extension of the laminar diffusion theory to turbulent flow is made by assuming first that the turbulent Schmidt number for concentration diffusion is unity, and secondly that pressure diffusion is not influenced by the turbulence.⁴ The turbulent viscosity is taken from the Keyes' correlation with the tangential Reynolds number of the vortex which is based on the diameter of the vortex tube. It is concluded that satisfactory containment is possible if certain limitations on the tangential Reynolds number are not exceeded.

The influence of the containment criteria on nuclear criticality and fission fragment absorption is considered next, and the overall performance of the system is estimated.

Though the analysis is carried out for a reactor consisting of a number of discrete jet driven vortex tubes embedded in a moderator material, the functional relations derived in the present study are applicable to other vortex reactor systems as well. Several of these schemes are discussed in terms of

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ Presently, Member of the Technical Staff, Aerospace Corp. Member ARS.

² Presently, Assistant Professor, Massachusetts Institute of Technology; and Consultant, Aerospace Corp. Member ARS.

³ Numbers in parentheses indicate References at end of paper.

⁴ A similar assumption has been made by Ragsdale (3).

their expected performance improvements as indicated by the analysis.

Containment Criteria

Diffusion in Turbulent Vortexes

A complete analysis of containment in laminar, inviscid vortexes has been given in (1). To adapt this analysis to turbulent flow, it is necessary to account for the influence of turbulence on the rates of diffusion. This can be done at present only in a semi-empirical way. There is, of course, a considerable body of experience justifying Reynolds' analogy, and its analog for mass transport, in turbulent pipe flows; however, in that case diffusion occurs only because of concentration gradients. In the vortex containment scheme, the containment is made possible by the relative diffusion of light and heavy molecular species in the strong pressure gradient of the vortex, i.e., by pressure diffusion. The influence of turbulence on the rates of pressure diffusion is, therefore, crucial to the present analysis.

It is proposed that the rate of pressure diffusion is in fact independent of the turbulent intensity, and exactly equal to that which would exist in a laminar flow with the same pressure gradient. It can be argued that this must be true, so long as the turbulent pressure fluctuations are small; however, the ultimate verification must be derived from comparison with experiments.

At present, the experimental data is very scanty, being restricted to one or two examples of gaseous separation in vortexes, as reported by Keyes and Dial (4). However, the agreement of Keyes and Dial's results with predictions based on the above postulate is sufficient to make a strong case for its validity.

As in (1), a two-dimensional vortex flow will be considered, the gas comprising a light species, typically hydrogen, and a heavy species, typically plutonium. It will be assumed that the two gases do not react chemically, so that each species satisfies a separate continuity equation. Alternatively, the overall continuity equation for the gas mixture, and that for heavy gas, may be used. If ρ is the gas density, u its radial velocity, and r the radius, the former equation is simply

$$-\rho ur = (Q_1 + Q_2)/2\pi \quad [1]$$

where Q_1 and Q_2 are the mass flows of light and heavy gas per unit of vortex length. Similarly, if ρ_2 and u_2 are the density and diffusion velocity of the heavy gas, its continuity equation is

$$-\rho_2(u + u_2)r = Q_2/2\pi \quad [2]$$

Assuming that the eddy diffusivity is ϵ times as large as the laminar diffusivity for concentration diffusion, and that the diffusivity is equal to the laminar value for pressure diffusion, the binary diffusion equation may be written (5)

$$u_2 \left(\frac{\rho}{\rho - \rho_2} \right) = -D_{12} \left[(1 + \epsilon) \frac{n}{n - n_2} \frac{d \ln (n_2/n)}{dr} + \frac{n(m_1 - m_2)}{\rho} \frac{d \ln p}{dr} \right] \quad [3]$$

where

- n, n_2 = total and heavy gas concentrations, respectively
- p = pressure
- m_1, m_2 = masses of the light and heavy molecules, respectively
- D_{12} = binary diffusion coefficient

The first term in the brackets represents the concentration diffusion, hence is multiplied by the factor $(1 + \epsilon)$, while the second term, representing the pressure diffusion, retains its laminar form.

It will be assumed that the turbulent Schmidt number is unity, i.e., that the turbulent diffusivity for mass transport is the same as the turbulent diffusivity for momentum transport. In this case, $(1 + \epsilon)$ equals the ratio of turbulent and laminar viscosities divided by the laminar Schmidt number, $Sc = \rho_1 D_{12} / \mu$.

It is convenient to rewrite the equation in terms of the density ratio, $w = m_2 n_2 / m_1 n_1$. Later arguments will show that w is not likely to exceed about 10 in practical cases, so that $n_2/n \ll 1$. With this fact, and the radial momentum equation, $dp/dr = \rho v^2/r$, Eq. 3 may be written in the form

$$1 - \left(\frac{1 + w}{w} \right) \left(\frac{Q_2}{Q_1 + Q_2} \right) = \frac{-2\pi \rho_1 D_{12}}{Q_1 + Q_2} \times \left[(1 + \epsilon) \frac{d \ln w}{d \ln r} - \left(\frac{m_2}{m_1} - 1 \right) \gamma M^2 \right] \quad [4]$$

where $M^2 = v^2/\gamma RT$ is the tangential Mach number based on the speed of sound in the light gas.

A complete solution of the diffusion problem requires the simultaneous solution of Eq. 4 with an energy equation and a tangential momentum equation, both of which enter into determining the variation of M^2 with r . Such solutions were obtained in (1), where it was found that under certain conditions, the density ratio w may have a simple maximum within the vortex, so that the fissionable material is in fact contained away from the periphery of the vortex, as indicated in Fig. 1.

From Eq. 4, it may be seen that, if such a simple maximum of w exists, then at the value of r corresponding to its position, the Mach number and mass flow rate are related by

$$\gamma M_m^2 = \frac{1}{[(m_2/m_1) - 1]} \left(\frac{Q_1 + Q_2}{2\pi \rho_1 D_{12}} \right)_m \times \left[1 - \left(\frac{1 + w_m}{w_m} \right) \frac{Q_2}{Q_1 + Q_2} \right] \quad [5]$$

where subscript m denotes quantities evaluated at the radial position of the maximum. This is, of course, exactly the same relation as was obtained for laminar diffusion, as there is no concentration diffusion at the stationary point of w . It follows that for a given radial mass flow rate and a given variation of M with r , the position of the maximum of w within the vortex is independent of the turbulence level. This conclusion is a direct consequence of the assumption that pressure diffusion is unaffected by turbulence, and its agreement with experiment, which will be described later, is therefore a justification of the assumption.

In view of the incomplete knowledge of the turbulent structure of vortexes, a detailed solution of the diffusion problem such as was obtained in (1) is not appropriate; rather a solution for w will be constructed by expansion about r_m . Thus, let

$$w(r) = w_m + \frac{1}{2} \left(\frac{d^2 w}{dr^2} \right)_m (r - r_m)^2 + \dots$$

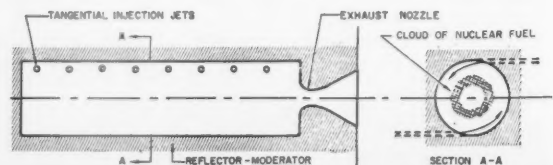


Fig. 1 Jet driven vortex tube

where the linear term is zero by definition of r_m . It will be convenient to introduce variables made dimensionless by division by their values at r_m , e.g., $r' = r/r_m$, $w' = w/w_m$.

The expansion will be carried only to the second order, as indicated. Assuming that $\rho_1 D_{12}$ is proportional to $(T')^\omega$, and differentiating Eq. 4 to obtain $(d^2w'/dr'^2)_m$, it is found that

$$w' = 1 + \frac{\gamma M_m^2 [(m_2/m_1) - 1]}{1 + \epsilon} \left[\left(\frac{dM'}{dr'} \right)_m + \frac{\omega}{2} \left(\frac{dT'}{dr'} \right)_m \right] (r' - 1)^2 \quad [6]$$

Since $M' = v'/T'^{1/2}$

$$\left(\frac{dM'}{dr'} \right)_m = \left(\frac{dv'}{dr'} \right)_m - \frac{1}{2} \left(\frac{dT'}{dr'} \right)_m$$

It will be demonstrated in the next section that for the cases of interest here, the velocity profile is approximately that of potential flow, i.e., $v' \approx (r')^{-1}$ so that

$$\left(\frac{dM'}{dr'} \right)_m \approx - \left[1 + \frac{1}{2} \left(\frac{dT'}{dr'} \right)_m \right] \quad [7]$$

In this case, Eq. 6 becomes

$$w' = 1 - \frac{\gamma M_m^2 [(m_2/m_1) - 1]}{1 + \epsilon} \left[1 + \frac{1}{2} (1 - \omega) \times \left(\frac{dT'}{dr'} \right)_m \right] (r' - 1)^2 \quad [8]$$

The value of ω is $\frac{1}{2}$ for "hard sphere" molecules, unity for Maxwellian molecules, and about $\frac{3}{4}$ for most gas pairs (5). From Eq. 8, it may be seen that, where $(dT'/dr')_m < -2/(1 - \omega)$, no peak of the density profile can exist. ω was taken as $\frac{1}{2}$ in (1) and it was concluded that $(dT'/dr')_m > -4$ was necessary. For $\omega = \frac{3}{4}$, this becomes $(dT'/dr')_m > -8$. The heating rate corresponding to these temperature derivatives was termed the "maximum heating rate" in (1). It is seen that the maximum heating rate for $\omega = \frac{3}{4}$ is twice that for $\omega = \frac{1}{2}$. An intermediate heating rate, or temperature derivative, exists beyond which the profile of w displays two stationary points, a maximum and a minimum. This was termed the "critical heating rate," and is an upper limit to the heating rate for a simple peaked distribution of w .

Corresponding to the critical heating rate there is a maximum overall temperature ratio compatible with the simple peaked distribution. Depending on M_m and w_m , this was found in (1) to be between two and four. There seems to be no simple way to estimate this critical heating rate for the present case, but it is quite likely that it increases as the maximum heating rate increases, so that the maximum overall temperature ratio increases as ω increases. It can be seen quite readily from Eq. 4 that for $\omega = 1$, there is no limit to the heating rate compatible with a simple peaked distribution of w .

In the following work, Eq. 8 will be assumed to describe the concentration profile of heavy gas in the vortex. This approximate profile actually lies below the profile given by integration of the complete system of differential equations. A typical profile from (1) is compared in Fig. 2 with the approximation to it given by Eq. 8. For this case, the approximate solution gives $w' = 0$ at a value of r' , which is 1.2 times the value for which the exact solution gives $w' = 10^{-4}$. Thus, a slight correction to the approximate theory is needed; the maximum permissible radius of the exit nozzle r_e' will be taken as 0.8 of r_c' , the value given by Eq. 8 for $w' = 0$.

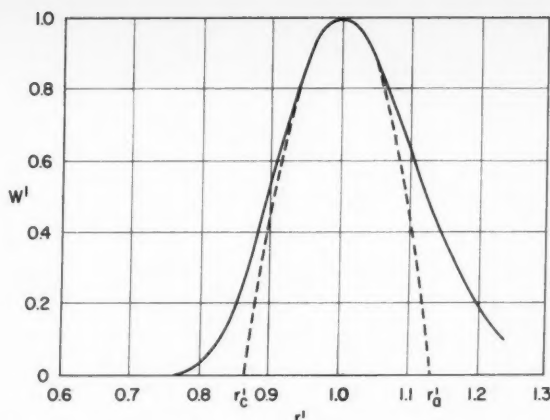


Fig. 2 Comparison of a typical concentration profile with its parabolic approximation: radius normalized with respect to position of maximum w'

Solving for r_c'

$$r_c' = 1 - \left\{ \frac{1 + \epsilon}{\gamma M_m^2 \left(\frac{m_2}{m_1} - 1 \right) \left[1 - \frac{(\omega - 1)}{2} \left(\frac{dT'}{dr'} \right)_m \right]} \right\}^{1/2} \quad [9]$$

Comparison With Separation Experiments

The two results of the above approximate analysis which can be compared with Keyes and Dial's experiment are: the position of the concentration peak, as given by Eq. 5, and the value of $(d^2w'/dr'^2)_m$, as implied by Eq. 8. As was mentioned above, the first comparison may be regarded as a test of the hypothesis that the rate of pressure diffusion is independent of the turbulence. Keyes and Dial have made this comparison for several cases [see (4), Table IV, p. 76] and find the agreement to be good. One of their cases has been analyzed in some detail⁶ [see (4), Case 3, Table IV, p. 76]. The Mach number profile, determined from the measured pressure profile, and the profile of w' are shown for this example in Fig. 3. For this case, $Q_1 = 0.0015$ lb per sec, and using Keyes' values of 0.27×10^{-3} ft² per sec at 80 F and 1 atm for D_{12} , and 0.5 for $(Q_2/Q_1)(w_m/w_c)$, Eq. 5 yields $M_m = 0.51$. The measured value is $M_m = 0.68$.

In the following work, it will be assumed that the turbulent Schmidt number is unity. To compare the prediction of $(d^2w'/dr'^2)_m$ based on this assumption with Keyes' measurements, it is necessary to determine the eddy viscosity from the measured pressure profile. This has been done in two ways. In the first, an analytical solution for the tangential velocity (see Eq. 11) was fitted to the Mach number profile of Fig. 3, assuming T constant. By this method, Re^* was found to be -0.67 , and $(dM'/dr')_m = -0.25$. These yield a profile of the form $w' = 1 - 0.17(r' - 1)^2$.

On the other hand, by fitting an analytical pressure profile, obtained by integration of the analytical velocity profile to the measured pressure, we find $Re^* = -1.8$. This gives $w' = 1 - 0.45(r' - 1)^2$.

The second method seems a more reliable one for deducing Re^* , so its prediction is compared with the data in Fig. 3. It underestimates the curvature somewhat, but considering the difficulty of estimating (dM'/dr') and Re^* , the agreement may

⁶ The authors are grateful to Dr. Keyes for personal communication of the pressure profile for Case 3.

⁷ See Eq. 11 for definition of Re^* .

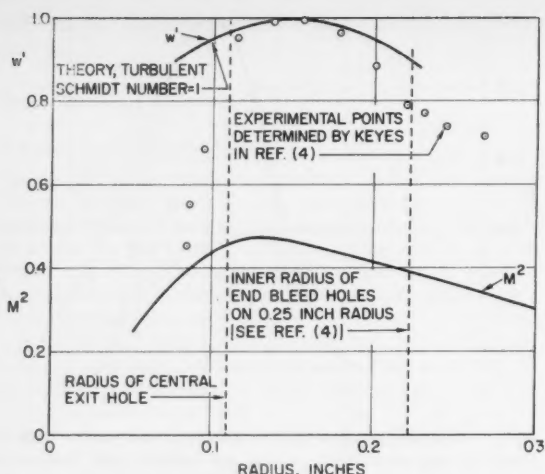


Fig. 3 Comparison of turbulent diffusion theory with experiment

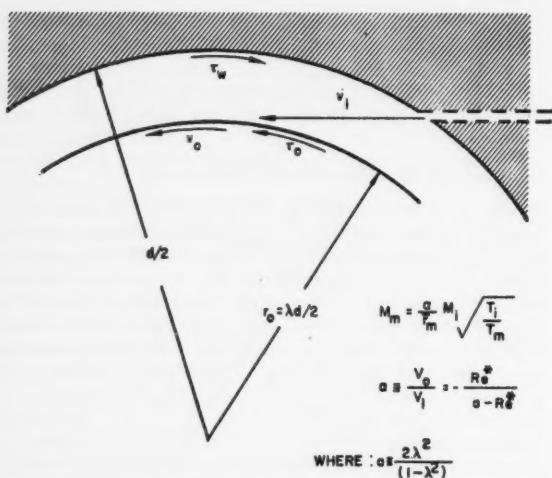


Fig. 4 Recovery of injection velocity

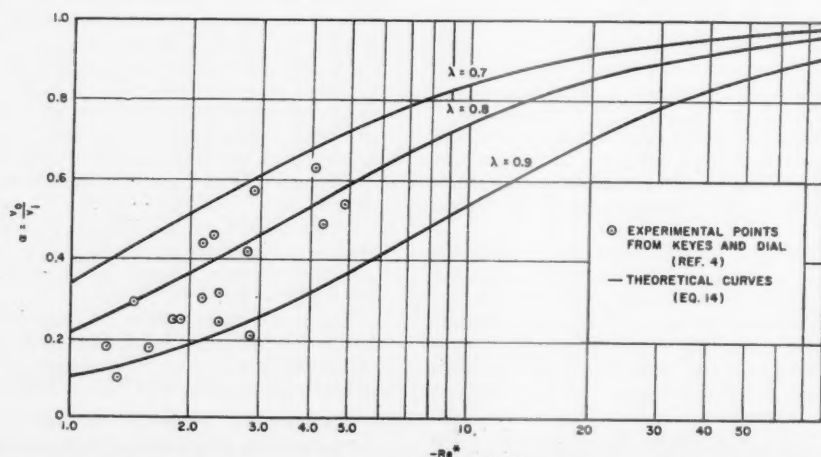


Fig. 5 Comparison of theoretical and experimental jet velocity recovery as a function of effective radial Reynolds number

be regarded as satisfactory. Clearly a more precise experiment is needed before the theory can be definitely confirmed or rejected.

In summary, Eq. 5 gives the relationship between the tangential Mach number at the peak of w , and the radial mass flow rate Q_1 . If the containment is satisfactory, $Q_2 \ll Q_1$, and since w_m is of order unity, the last term in the brackets is negligible. Dropping it and writing Q_1 in terms of the radial Reynolds number $Re = \rho u r / \mu = -Q_1 / 2\pi\mu$, we find

$$M_m^2 \approx \frac{-Re}{\gamma Sc [(m_2/m_1) - 1]} \quad [10]$$

In the following calculations, it will be assumed that $\gamma = 1.3$, $Sc = 1$, $m_2/m_1 = 120$, so that $M_m^2 = -0.0065 Re$. This expression and Eq. 9 complete the description of the separation process for the present purpose. In the next section, the specification of ϵ and M_m , in terms of fluid mechanical parameters of the vortex flow, will be described.

Determination of Fluid Mechanical Relations

In order to arrive at a criterion for satisfactory containment it is necessary to relate the turbulent diffusivity to the fluid mechanical variables, and to determine how the velocity varies through the vortex.

Determination of the velocity profile is relatively simple. Since heat addition and compressibility do not affect the tangential momentum equation in two-dimensional flow, a straightforward approach such as that used by Einstein and Li (6) can be employed to find the velocity profiles. Outside the exit hole radius, (6) gives as a solution

$$\frac{v}{v_0} = A \left(\frac{r}{r_0} \right)^{-1} + (1-A) \left(\frac{r}{r_0} \right)^{Re^*+1}, \quad r > r_0 \quad [11]$$

where

$$A = \frac{Re^* + 2(1 - e^{Re^*/2})}{Re^* \{1 - e^{Re^*/2} [(r_e/r_0)]^{-(Re^*+2)}\} + 2(1 - e^{Re^*/2})}$$

and Re^* is the radial Reynolds number based on the turbulent viscosity (laminar plus eddy) and has been taken as constant.

For any fixed value of (r_e/r_0) between zero and one, A approaches unity exponentially as $(-Re^*)$ becomes infinite. In fact, for $r \geq r_m$ the approximation

$$\frac{v}{v_0} = \left(\frac{r}{r_0} \right)^{-1} \quad [12]$$

is accurate to within 15% as long as $(-Re^*) > 3$ and $(r_e/r_0) > 0.1$. Since this assumption is quite accurate and results in considerable algebraic simplification, Eq. 12 will be adopted throughout the remainder of the analysis.

It is necessary to determine the actual peripheral velocity in the vortex in terms of a physically measurable parameter. For jet driven vortexes, the obvious choice is the tangential injection velocity. The fraction of the jet velocity that is recovered in the vortex may be found by a balance between the difference of the angular momentum fluxes in the jets and exit flow and the retarding torque due to shear at the tube wall. Such a calculation was carried out by Rosenzweig (7) for the case of laminar flow. The geometry considered is shown in Fig. 4. Using the velocity profiles given by Eq. 11 the calculation results in the following expression for jet recovery

$$\alpha \equiv \frac{v_0}{v_i} = \left\{ 1 - \frac{1}{Re} \left[\frac{2}{1 - \lambda^2} - 2A + Re(1 - A) \right] \right\}^{-1} \quad [13]$$

where λ is the ratio of injection radius (r_0) to the total radius of the vortex tube ($d/2$).

Consistent with the approximation of Eq. 12 (i.e., $A = 1$), Eq. 13 may be rewritten in the form

$$\alpha = -Re/(a - Re) \quad [14]$$

where

$$a \equiv 2\lambda^2/(1 - \lambda^2)$$

A similar result was obtained by Kerrebroek and Meghreblian (1) who performed an approximate torque balance (and erroneously omitted a factor of 2).

Although the calculation in (7) was performed for laminar flow, experimental evidence indicates that Eq. 14 is at least approximately correct for turbulent flow if the laminar Reynolds number is replaced by an appropriate turbulent value based on the total viscosity, i.e., Re^* .

Eq. 14 is plotted in Fig. 5 for several values of λ . Experimental points from (2 and 4) are also shown in Fig. 5. The geometrical values of λ for these experiments ranged from about 0.88 to 0.95. It is clear from the comparison that the effective value of λ is somewhat smaller than the geometrical value. Two phenomena which could contribute to this difference are the mixing of jets and the turning which a supersonic jet suffers if the exit plane of the nozzle is not perpendicular to the nozzle axis. Within the limit of this uncertainty, the data indicate the validity of Eq. 14.

Keyes (2) has experimentally determined the variation of turbulent eddy viscosity in jet-driven vortex tubes. He found that the ratio of eddy viscosity to molecular viscosity correlated with the tangential Reynolds number evaluated at the periphery of the vortex. The experimental data scatter about a least squares line given by the expression (see Fig. 6)

$$\mu^*/\mu = C_1(Re_{t0})^k \quad [15]$$

where

$$\begin{aligned} Re_{t0} &= v_0 \rho_0 d / \mu_0 \\ C_1 &= 2.03 \times 10^{-3} \\ k &= 0.86 \end{aligned}$$

The uncertainty in these data is rather large. Keyes determined μ^*/μ by comparing velocity profiles deduced from experimentally determined pressure distributions, with analytical profiles, at only one radius in the vortex. In reality the eddy viscosity may be expected to vary throughout the vortex and thus the results would depend on the point where the comparison was made. Nevertheless, these are the only data of this sort available at the present time.

If, however, the functional form of Eq. 15 is correct (and there seems to be little reason to doubt its validity) the indi-

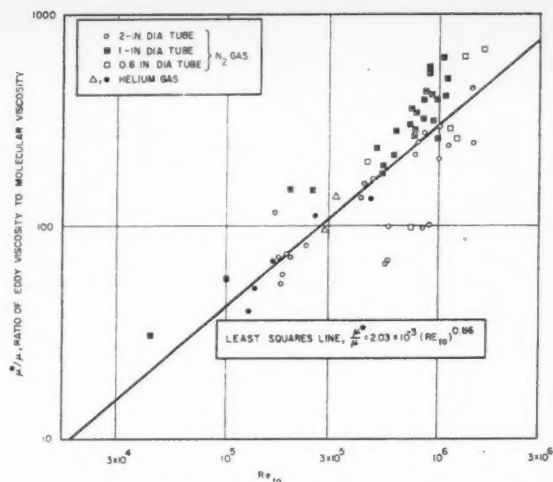


Fig. 6 Variation of eddy viscosity with tangential Reynolds number [from (2)]

cation is that the turbulent level in the vortex is determined by the boundary layer on the outer cylinder and not directly by the jets, as some had previously conjectured. Furthermore, if one assumes that the boundary layer at the outer cylinder is similar to a flat plate boundary layer at the same Reynolds number (such that the shear at the wall is proportional to $Re^{-0.2}$), then the exponent k turns out to be 0.8 (compared to 0.86 for Keyes' experiments).

Results for Satisfactory Containment

It is now possible to combine the previous relations and thus define the regions of fluid mechanical variables for which satisfactory containment of a fissionable gas, in a turbulent vortex, is possible. Using the theoretical relation between mass flow and Mach number from the diffusion analysis (Eq. 10) together with the experimental results from Keyes (Eq. 15) gives

$$\begin{aligned} M_m^2 &= -C_2 Re = -C_2 (\mu^*/\mu) Re^* \\ &= -C_1 C_2 Re^* (Re_{t0})^k \end{aligned}$$

or

$$Re_{t0} = \left(-\frac{M_m^2}{C_1 C_2 Re^*} \right)^{1/k} \quad [16]$$

Also, from Eqs. 12 and 13

$$M_m^2 = \frac{\alpha^2}{\bar{r}_m^2} M_i^2 \frac{T_i}{T_m}$$

where a bar over a symbol indicates it is normalized with respect to its value at $r = r_0$.

The radius of the maximum concentration \bar{r}_m is related to \bar{r}_e and Re^* (through Eqs. 9 and 10) by the formula

$$\bar{r}_m = \frac{[-Re^*(1+b)]^{1/2}}{[-Re^*(1+b)]^{1/2} - 1} \bar{r}_e \quad [17]$$

where

$$b \equiv \frac{1 - \omega}{2} \left(\frac{dT''}{dr'} \right)_m$$

Using the above expression for \bar{r}_m , and substituting for α from Eq. 14 gives

$$M_m^2 = - \frac{M_i^2}{\bar{r}_e^2} \left(\frac{T_i}{T_m} \right) \frac{Re^*}{(a - Re^*)^2} \times \frac{\{[-Re^*(1+b)]^{1/2} - 1\}^2}{(1+b)} \quad [18]$$

Eq. 16, with M_m given by Eq. 18, is plotted in Fig. 7 for a few values of (T_i/T_m) and b . The curves are lines of constant \bar{r}_e . Thus, the maximum value of \bar{r}_e for satisfactory containment is defined for each point in the plane. Note that for any given value of \bar{r}_e , there is a maximum value of Re_{e0} for which the vortex will satisfactorily contain the fissionable material. This limit represents a rather strong restriction on the fluid mechanical variables, as will be seen in the next section when criticality requirements are considered.

The curves in Fig. 7 are cut off at $Re^* = -3$, since below this point the approximation of Eq. 12 is no longer valid. Also, it is rather dangerous to extrapolate Keyes' experimental results (Fig. 6) to values of Re_{e0} which correspond to very low values of μ^*/μ .

Nuclear Considerations

Criticality Requirements

In the previous section a formula was derived which details the fluid mechanical requirements for the successful operation of a vortex as a containment device. If the reactor is to generate power, it must contain fissionable material in sufficient concentration to support a chain reaction. This criticality requirement is a function of the reactor geometry and composition, and may be characterized by the macroscopic fission cross section (Σ_f) for the reactor. For a specific geometrical configuration and composition, the value of Σ_f determines the size and weight of the reactor. As Σ_f decreases, the size increases, until, for some value of Σ_f , an infinite reactor is required for criticality; beyond this value it is impossible to sustain a nuclear reaction.

The vortex reactor envisioned in the present study consists of a large number of small vortex tubes interspersed in a neutron-moderating material and surrounded by a beryllium reflector. The reactor is assumed to be a right circular cylinder with equal height and diameter. Criticality calculations have

been carried out for such systems (8), in which graphite and beryllium oxide were considered as moderators. It was found that there were optimum values of the reflector thickness and void fraction (δ) which minimized the reactor weight. These values were about 15 cm and 0.6 respectively. The mass of such reactors is given in Fig. 8 as a function of the macroscopic fission cross section, averaged over the core.

When all the fissionable material is in gaseous form, the macroscopic fission cross section for the reactor can be written

$$\Sigma_{fg} = \delta n_f \sigma_f$$

where n_f is the average concentration of fissionable material in the vortex tubes and σ_f is the microscopic fission cross section.

If some of the fissionable material in the reactor is contained in the solid phase (e.g., distributed throughout the moderator) then the overall macroscopic fission cross section becomes

$$\Sigma_f = \frac{\delta n_f \sigma_f}{1 - f} \quad [19]$$

where f denotes the fraction of the fissionable material in the solid.

The average concentration n_f can be written in terms of the fluid mechanical variables as follows

$$\begin{aligned} n_f &= \frac{\rho_{1 \text{ avg}} w_m w'_{\text{avg}}}{m_2} \\ &= \frac{\rho_s}{m_2} \left(\frac{\rho_{1 \text{ avg}}}{\rho_s} w_m w'_{\text{avg}} \right) = \frac{\rho_s}{m_2} \omega \end{aligned} \quad [20]$$

where the subscript s denotes standard conditions in the light-weight gas (e.g., hydrogen), and

$$\omega = \frac{\rho_{1 \text{ avg}}}{\rho_s} w_m w'_{\text{avg}} \quad [21]$$

The parameter ω represents a dimensionless fuel concentration in the vortex, and is a nuclear figure of merit for the overall system.

Taking

$$\begin{aligned} \delta &= 0.6 \\ \sigma_f &= 0.9 \times 10^{-21} \text{ cm}^2 \\ m_2 &= 4 \times 10^{-22} \text{ g} \\ \rho_s &= 0.82 \times 10^{-4} \text{ g cm}^{-3} \end{aligned}$$

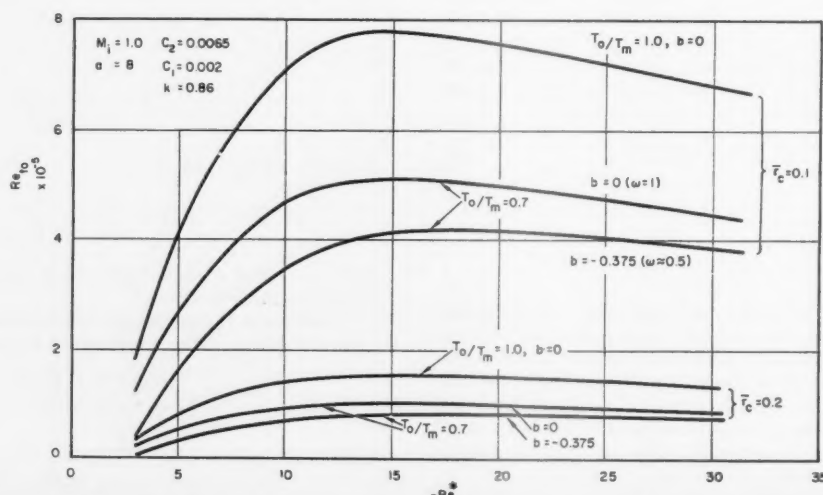


Fig. 7 Maximum tangential Reynolds number for satisfactory containment, as a function of effective radial Reynolds number

the final result for Σ_f is

$$\Sigma_f = 1.9 \times 10^{-4} \frac{\omega}{1-f} \text{ cm}^{-1} \quad [22]$$

From Fig. 8, the minimum practical value of Σ_f is about $0.3 \times 10^{-3} \text{ cm}^{-1}$, so that $\omega/(1-f)$ must be greater than 1.5 if a plutonium-hydrogen system as described is to be critical.

Derivation of Attainable Fuel Concentration

In this section a rough estimate will be made of the values of ω that are attainable within the restrictions imposed on the system by the requirements for containment.

Examining the parameter ω in more detail, it is seen that Eq. 21 can be rewritten

$$\omega = \frac{\rho_{1 \text{ avg}}}{\rho_s} w_m w'_{\text{avg}} = \frac{\rho_{1 \text{ avg}}}{\rho_0} w_m w'_{\text{avg}} \times$$

$$\frac{Re_{t0}}{\alpha M_i} \left(\frac{\mu_0}{d \rho_s \sqrt{\gamma R T_0}} \right) \sqrt{\frac{T_0}{T_i}} \quad [23]$$

As long as M_i is of order one, $T_0/T_i \approx 1$, so that the last factor in Eq. 23 may be ignored in the present analysis. Further, if it is assumed that $\mu \sim T^{1/2}$ then the grouping $(\mu_0/\rho_s \sqrt{\gamma R T_0})$ is a constant property of the light gas and is independent of temperature. For hydrogen this constant is approximately: $9 \times 10^{-6} \text{ cm}$.

The quantity in parentheses on the right hand side of Eq. 23 is the inverse of a characteristic Reynolds number based on the properties of the light gas and the diameter of the vortex tubes. Since the maximum value of Re_n is limited by the turbulent containment criteria, it is evident that $\omega \sim 1/d$, indicating that small diameter tubes are a requirement for criticality.

Proceeding to the evaluation of the other terms in the expression for ω , the function w'_{avg} can be found by integrating the approximate parabolic concentration profile over the vortex. Defining the outer radius of the parabolic concentration profile by \bar{r}_a (with $\bar{r}_a = 2\bar{r}_m - \bar{r}_c$), (see Fig. 2) the integration yields

$$w'_{\text{avg}} = \frac{8}{3} \bar{r}_c^2 \frac{[-Re^*(1+b)]^{1/2}}{\{[-Re^*(1+b)]^{1/2} - 1\}^2} \quad \bar{r}_a < 1 \quad [24]$$

Evaluation of $\rho_{1 \text{ avg}}/\rho_0$ in terms of other variables is not quite as simple. The difficulties are reduced somewhat if it is assumed that $\rho_{1 \text{ avg}} \approx \rho_{1 m}$. This assumption is valid as long as $\bar{r}_m \approx 0.5$. Estimation of $\rho_{1 m}/\rho_0$ is accomplished by integration of an approximate radial momentum equation.

When $(u/v)^2 \ll 1$, the radial momentum equation reduces to

$$\frac{d \ln p}{d \ln r} = \gamma M^2 \quad [25]$$

This equation can be integrated by dividing the vortex into two regions and making some further simplifying assumptions. In the region outside of the annulus containing the nuclear fuel, the velocity is proportional to $1/r$ while the temperature remains very nearly constant. Therefore, when $\bar{r} > \bar{r}_a$

$$\frac{d \ln p}{d \ln \bar{r}} = \gamma \frac{M_0^2}{\bar{r}^2}, \quad \bar{r} > \bar{r}_a \quad [26]$$

Within the annulus bounded by \bar{r}_c and \bar{r}_a , it is assumed that the Mach number is constant and equal to its value at the point of maximum concentration; thus, in this region

$$\frac{d \ln p}{d \ln \bar{r}} = \gamma M_m^2 \left(1 + \frac{2}{3} w_m \right) \quad [27]$$

The term in parentheses in Eq. 27 represents a density correction due to the presence of the heavy fuel concentration in this

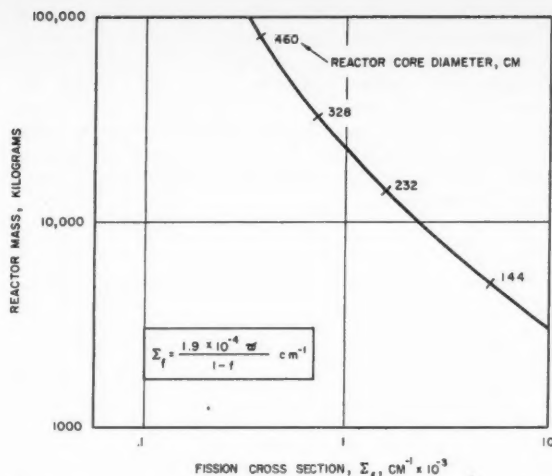


Fig. 8 Mass and radius of beryllium reflected graphite reactors with length and diameter equal—reflector thickness, 15 cm.

region, $[(2/3)w_m]$ being the average value of w over the annulus.

Using Eqs. 26 and 27, the value of p_m can be found by integration. The result, then, for $\rho_{1 m}/\rho_0$ is

$$\frac{\rho_{1 m}}{\rho_0} = \frac{T_0}{T_m} \frac{p_m}{p_0} = \frac{T_0}{T_m} \exp \left[-\frac{\gamma}{2} \alpha^2 M_i^2 (\bar{r}_a^{-2} - 1) + \gamma M_m^2 \left(1 + \frac{2}{3} w_m \right) \ln \left(\frac{\bar{r}_m}{\bar{r}_a} \right) \right] \quad [28]$$

where M_m is given by Eq. 17.

Thus far, the parameter w_m has been left free. Since a large value of ω is obviously desirable from the point of view of reactor size, it might seem at first glance that w_m should be made as large as possible. However, w_m appears in the exponent in the expression for ρ_m/ρ_0 , and so cannot become very large before ω becomes minute. The parameter w_m can be eliminated from further consideration by optimizing ω with respect to it. This procedure leads to the result

$$(w_m)_{\text{max}} \omega = \frac{1}{(2/3) \gamma M_m^2 \ln (\bar{r}_a/\bar{r}_m)} \quad [29]$$

Putting these expressions back into ω (and using the results previously given to eliminate Re_{t0}) provides a formula for ω in terms of the seven remaining variables, (T_0/T_m) , M_i , a , b , \bar{r}_c , d and Re^* . The resulting curves are plotted in Fig. 9, for typical values of the various parameters, for a system using hydrogen and plutonium gas.

Note that the largest values of ω occur at the lower values of $(-Re^*)$, and that even these values correspond to values of Σ_f too small for criticality. The low values of ω are a direct result of the turbulence limitation relating Re_{t0} to Re^* (Eq. 16). The fact that the tangential Reynolds number must be small to prevent the turbulence from broadening the concentration profile of the heavy gas implies that the peripheral pressure must also be relatively low. This results in the vortex being unable to contain a sufficient concentration of fissionable material for criticality. With the inviscid approximation used here, ω increases continuously as $-Re^*$ decreases. If the deviation of the velocity profile from $1/r$ were included at low values of $-Re^*$, ω would decrease to zero at $Re^* = -1$. The broadening of the concentration profile is the most serious manifestation of the turbulence in the vortex, for even if full recovery were achieved (i.e., $\alpha = 1$) the restriction implied by Eq. 16

would still be present. This difficulty can be circumvented by placing some of the fuel in the solid phase in the moderator (i.e., making $f > 0$) and thus raising Σ_f according to Eq. 19. Of course, such a scheme cannot be carried too far since, as f increases, radiators are required to remove the excess energy generated in the solid, and the reduction in the reactor weight achieved by raising Σ_f may be offset by increased radiator weight. This problem is considered in detail in the section on overall performance.

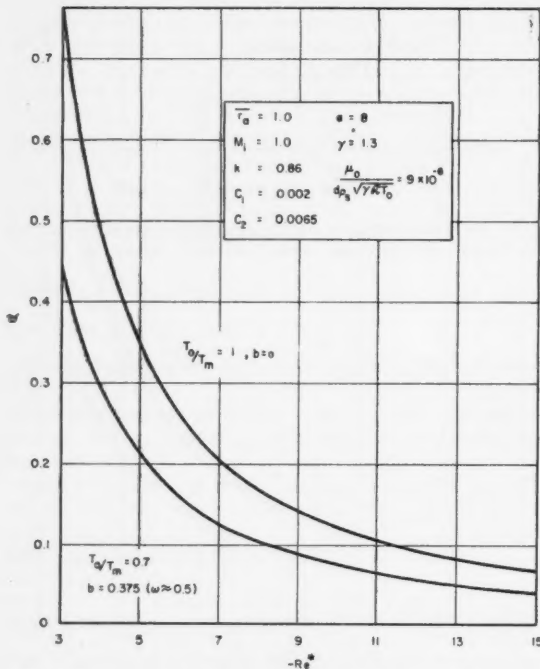


Fig. 9 Concentration of fissionable material as a function of effective radial Reynolds number

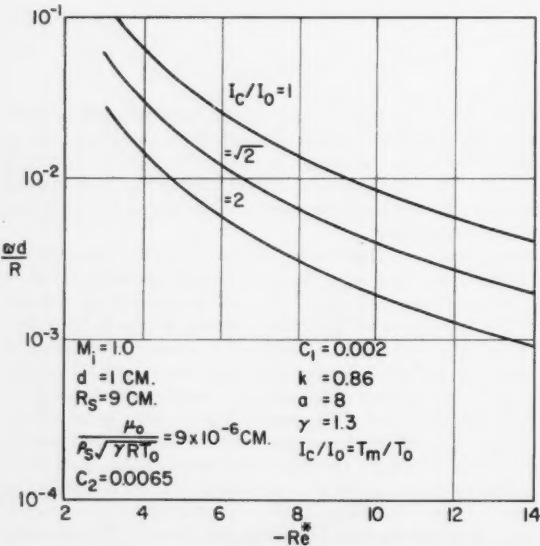


Fig. 10 Concentration of fissionable material divided by the range of fission fragments as a function of effective radial Reynolds number

Fission Fragment Absorption

The combination of small vortex tubes and low pressures which results from the containment requirement raises the question whether the kinetic energy of the fission fragments generated in the gas will really be deposited there. If the range of the fission fragments is large compared to the tube diameter, then a large fraction of their energy will be deposited in the surrounding solid material.

At standard density, the range of fission fragments in hydrogen is about 9 cm (9), and it is inversely proportional to the electronic concentration of the scattering material. The ratio of the total electronic concentration in the vortex to that of the hydrogen is $1 + 0.39 w$, so that if R is the fission fragment range, and ρ_1 is the density of the hydrogen

$$\frac{R}{R_s} = \frac{\rho_s}{\rho_1(1 + 0.39 w)}$$

Clearly the ratio of R to the tube diameter d is a measure of the fraction of the fission fragments which impinge on the tube walls. It can be written in terms of the tangential Reynolds number Re_{t0} as

$$\frac{R}{d} = \frac{R_s \rho_s \sqrt{\gamma R T_0}}{\mu_0} \frac{\alpha M_1}{Re_{t0}} \quad [30]$$

if the local value of R at the periphery is assumed to be the value everywhere within the vortex. Typical values of αM_1 , $(R_s \rho_s \sqrt{\gamma R T_0} / \mu_0)$, and Re_{t0} , as derived from the containment analysis, are about 0.3, 10^6 and 10^4 , so that R/d will be of the order of 30.

For such large values of R/d , it is readily possible to estimate the fraction of the fission energy which is deposited in the tube walls. It will be assumed first that the tubes may be replaced by a plane-parallel configuration in which two concentration peaks of fissionable material are symmetrically disposed in the space between plane-parallel walls. Secondly, it will be assumed that the energy of the fission fragments decays linearly from the point of their origin, with a rate equal to the actual rate at the point of origin. The rate of decay, or the energy deposition rate, in fact decreases rapidly with the distance (9); however, since only a small fraction of the path will be within the tubes, the present approximation should be reasonable. From the data of (9), the initial rate of energy decay is $d(E/E_0)/d(r/R) = -2.5$, where E_0 is the initial energy, and r is the distance from the point of fission. With the assumed linear decay, the plane parallel geometry, and the symmetry about a plane midway between the absorbing surfaces, it is clear that the actual symmetrical fission distribution may be replaced by a concentration in the central plane. The attenuation distance then becomes one half the separation of the surfaces, which is to be taken here as the tube radius. Thus, if ξ is the fraction of the energy of the fission fragments which is deposited in the walls

$$\xi = 1 - 1.25(d/R) \quad [31]$$

This formula is, of course, only valid for $R/d \gg 1$.

Reactor Performance

From the viewpoint of overall performance, the factors which are of importance are the total mass of the propulsion system, its thrust to weight ratio and its specific impulse.

The total mass will consist largely of two parts, the mass of the reactor, and that of the radiators. Criticality requirements determine the size, and therefore the mass of the reactor, as given in Fig. 8.

On the other hand, the mass of the radiator depends on the amount of heat which must be rejected from the solid parts of the reactor; it is naturally desirable to minimize it. Let ξ denote the ratio of the radiated power to the power actually delivered to the gas, by fuel in the gas phase; a simple heat

balance then shows that

$$\xi = \frac{1}{(1-f)(1-\zeta)} - \frac{(I_c/I_0)^2}{(I_c/I_0)^2 - 1} \quad [32]$$

where I_c is the final specific impulse of the device, and I_0 is the specific impulse which it would have if the propellant were exhausted at the enthalpy level to which it is raised by heating in the solid parts of the reactor. Now, if C_R is the mass of the radiator per unit of radiated power, it follows that the radiator mass \mathcal{M}_{rad} is

$$\mathcal{M}_{rad} = \frac{g^2}{2} C_R \xi F I_0 \left(\frac{I_0}{I_c} \right) \left[\left(\frac{I_c}{I_0} \right)^2 - 1 \right] \quad [33]$$

where F is the thrust.

Since the mass flow rate per unit of vortex tube length is essentially invariant, the thrust depends only on the total length of tube in the reactor, and I_c . As the tubes are made smaller, the thrust increases. A lower limit to the tube size is set by the requirement that the mass flow leave the tube at one end, through a nozzle which is smaller in radius than $0.8r_c$, the corrected inner radius of the fissionable gas distribution. This criterion was used in (1), but is difficult to apply in the present approximate analysis because of the difficulty of computing the radial pressure drop precisely. In lieu of it, a value will be assumed for the length to diameter ratio of the vortex tubes. If the tubes are assumed to extend through the length of the reactor, this ratio becomes D/d . Now the total length of tube in the reactor is $\delta D^2/d^2$, and the thrust is therefore

$$F = -2\pi\mu Re\delta D(D/d)^2 I_c \quad [34]$$

Taking the overall thrust to weight ratio as $F/W = F/(\mathcal{M}_{react} + \mathcal{M}_{rad})$

$$\frac{F}{W} =$$

$$\frac{1}{\left\{ \frac{-\mathcal{M}_{react}}{2\pi\mu Re\delta D I_0} \left(\frac{d}{D} \right)^2 \left(\frac{I_0}{I_c} \right) + \frac{g^2}{2} C_R I_0 \left(\frac{I_0}{I_c} \right) \xi \left[\left(\frac{I_c}{I_0} \right)^2 - 1 \right] \right\}} \quad \frac{\text{lb thrust}}{\text{lb mass}} \quad [35]$$

where ξ is given by Eq. 32. Taking the value of ζ from Eq. 31 and f from Eq. 19, ξ becomes

$$\xi = 4.2 \times 10^3 \Sigma_f \left(\frac{R}{d\omega} \right) - \frac{(I_c/I_0)^2}{(I_c/I_0)^2 - 1} \quad [36]$$

with Σ_f in cm^{-1} .

The parameter $\omega d/R$, which expresses the influence of both the criticality requirement and the fission fragment absorption on the required heat rejection from the solids, is shown in Fig. 10. It has been computed from Eqs. 23 and 30 by maximizing with respect to r_c , and taking $T_0/T_m = I_0/I_c$. The result is

$$\frac{\omega d}{R} = 4 \frac{d}{R_s} \left(\frac{\mu_0}{d\rho_s \sqrt{\gamma R T_0}} \right)^2 \left(\frac{I_0}{I_c} \right)^{2/k} \times \frac{\gamma^{1-2/k} \exp \{ [\gamma Re^* M_i^2 / 2(a - Re^*)^2] + 1 - 2/k \}}{(-C_1 C_2 Re^*) 2k \beta \ln [(\beta + 1)/\beta] (4/k - 4)^{2-2/k}} \times \left[\left(\frac{\beta}{\beta + 1} \right)^2 + 2 \sqrt{\frac{I_0}{I_c}} \ln \frac{\beta + 1}{\beta} \right]^{2-2/k}$$

with

$$\beta = [-Re^*(1+b)]^{1/2}$$

The curves were found to be rather insensitive to variation in b ; therefore this variation is not shown in Fig. 10.

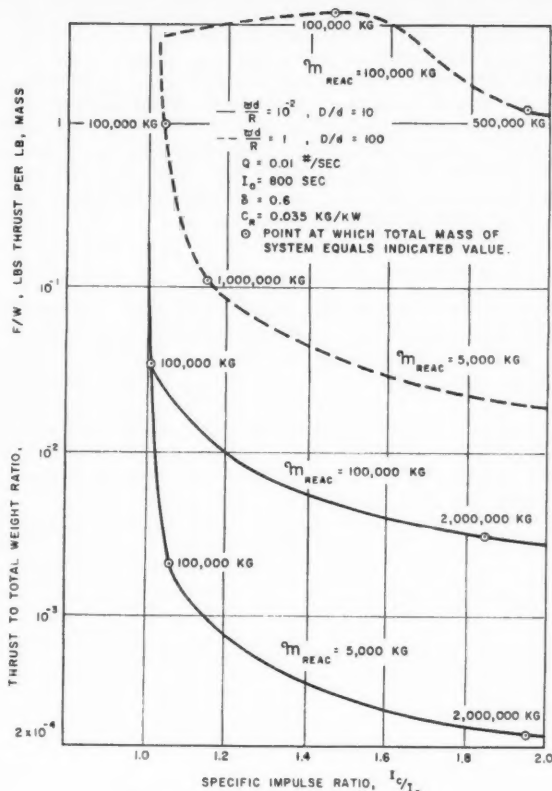


Fig. 11 Vortex reactor performance

The first term in the denominator of Eq. 35 is the weight-to-thrust ratio of the reactor itself. For $D/d = 10$, $\delta = 0.6$, $I_0 = 800$ sec, and $2\pi\mu Re = -0.01$ lb/sec-ft, it is 4.9 for a 5000 kg reactor and 28 for a 100,000 kg reactor. The magnitude of the second term of course depends crucially on C_R ; it will be assumed that the radiator mass per unit area is 10 kg m^{-2} , and that the temperature is 1500 K . C_R then becomes 0.035 kg/kw , and the factor $g^2 C_R I_0 / 2$ is 1.35. The overall thrust-to-weight ratio is given in Fig. 11 for 5000 and 10,000 kg reactors, as a function of I_c/I_0 . The solid curves correspond to $\omega d/R = 10^{-2}$, and $D/d = 10$, which, from Fig. 10, are characteristic of the values attainable with single turbulent vortices. The thrust to weight ratios are rather low, but even more important are the extremely large total masses, which are indicated along the curves.

If the relationship between the turbulence level and the tangential Reynolds number can be altered, the performance will improve markedly. This is indicated by the dashed curves, for which it has been assumed that the peripheral pressure in the vortex is 10 times that for the solid curves while the turbulence level remains the same. This results in an increase of $\omega d/R$ to one, and D/d to 100. The thrust to weight ratio for the 100,000 kg reactor is quite high for this case, because the radiator weight is small; however, the total mass is still very large.

The total mass can be reduced further by increasing the pressure even more, or by reducing the fission fragment absorption in solids. Some possible methods for effecting such improvements are discussed in the next section.

Advanced Techniques for Improving Reactor Performance

Although the numerical calculations of the preceding section are admittedly approximate, it is felt that the functional relationships derived in the paper are essentially correct. Most of these relationships are applicable to other vortex reactor systems in addition to the one described thus far, and so may be used to predict, in a qualitative way, what performance improvements may be expected from the various modifications of the basic vortex reactor concept. In this section several of these variations are discussed including the most promising of these, the vortex matrix.

Rotating Wall System

One system that has been proposed to reduce the turbulence in vortices is the rotating porous wall. In this device, the jets are eliminated, and the vortex is driven by a rotating porous wall through which the propellant is introduced. Such a scheme could conceivably reduce the turbulence level since the boundary layer on the outer wall is effectively eliminated. Kendall (10) has reported measuring lower turbulence levels in a rotating wall system as compared to a jet driven system, although no detailed data have yet been published.

If the rotating wall idea worked to perfection and generated truly laminar vortices, the restrictions on the gas density and pressure would be removed, thus relieving the criticality and fission fragment absorption problems; however, the mass flow per unit of vortex tube length remains limited by the diffusion process, so that a large number of small tubes would still be required to produce significant thrust levels. Also, the wall velocity must be a substantial fraction of the sonic velocity in hydrogen at 2500 K. The combination of high speed, high temperature, and the large number of tubes seem to pose a rather severe mechanical problem.

Recirculating Systems

A second scheme for generating strong vortices employs a large amount of mass injection in the form of jets to drive the vortex, whereupon most of the injected mass is bled off axially at some radius and is recirculated. The bleed to injection ratio is adjusted so that the mass flow that passes through the remainder of the vortex satisfies the containment criterion. In effect, this system creates an artificial rotating wall at some radius less than the tube radius. It was actually employed by Keyes and Dial (4) in the separation experiments performed at Oak Ridge National Laboratory.

The problems associated with the recirculation and cooling of large amounts of hot hydrogen may preclude the practical application of this idea to a reactor. However, studies of this system should be continued. In particular, the influence of the bleed off on the turbulence level in the inner portion of the vortex should be studied.

Vortex Matrix System

Perhaps the most novel and most promising of the advanced techniques involves the generation of a large number of stable vortices in a single cavity (11). The so-called vortex matrix (see Fig. 12) consists of a large number of small injection tubes arranged in a square array and perforated so as to inject fluid into the cavity in the directions shown in Fig. 12. Exit holes are located in the end walls at the centers of the squares formed by the injection tubes. As can be seen in Fig. 12 the adjacent vortices "roll" on each other; the pattern can be extended indefinitely. As the number of vortex cells increases, the wall area per vortex decreases, thus reducing the net momentum loss per vortex. By elimination of physical boundaries between vortices it is hoped that this system will improve jet recovery and reduce the overall turbulence level of the system. A pilot model of this device, having nine cells

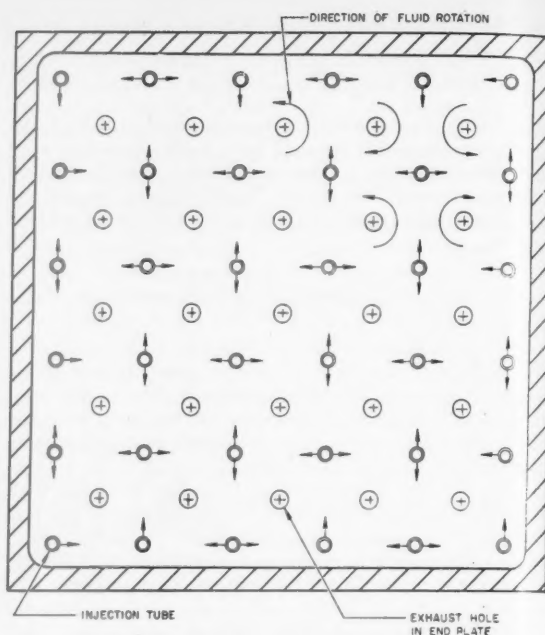


Fig. 12 Cross-section of vortex matrix

and using water as the working fluid, has been built and successfully tested at Space Technology Laboratories. More recently, a 25-cell matrix, using nitrogen, has been operated successfully at Aerospace Corporation. Much work remains in the actual determination of velocity profiles and turbulence levels.

Further work on this novel idea seems called for.

Concluding Remarks

The applicability of a turbulent, jet driven vortex as a containment device for the gaseous fission rocket has been studied. It has been concluded that containment is possible but only at low pressures because of the broadening effects of turbulence on the concentration profile of the fissionable gas. This restriction requires that some fissionable material be located in the solids in order for the reactor to be critical. In addition, a significant fraction of the energy generated in the gas is deposited in the solids, due to the relatively large fission fragment range. These effects require the use of radiators to reject heat from the solid parts of the reactor. Generally, the radiator weight dominates the reactor weight and this fact is reflected in the rather low thrust-to-weight ratios attainable for the overall system.

It has been pointed out that improvements on the basic vortex scheme must be aimed at altering the dependence of the turbulence level on the pressure so that higher pressures and reduced turbulence levels can be maintained simultaneously. Only in this way can substantial improvement in thrust to weight ratio be achieved.

Nomenclature

- A = function defined in Eq. 11
- a = function defined in Eq. 14
- b = function defined in Eq. 17
- C_R = radiator coefficient
- D = reactor diameter
- D_{12} = binary diffusion coefficient
- d = diameter of vortex tube

F = thrust
 f = fraction of fissions in the solid
 g = acceleration of gravity
 I = specific impulse
 k = exponent defined in Eq. 15
 M = Mach number
 \mathcal{M} = mass
 m = molecular weight
 n = concentration
 p = pressure
 \dot{Q} = mass flow
 R = universal gas constant
 R = fission fragment range
 Re = radial Reynolds number
 Re_t = tangential Reynolds number
 r = radial coordinate
 Sc = Schmidt number
 T = static temperature
 u = radial velocity
 v = tangential velocity
 W = weight
 w = ratio of density of heavy gas to light gas
 α = function defined in Eq. 14
 γ = ratio of specific heats of light gas
 δ = void fraction of reactor
 ϵ = ratio of turbulent diffusivity to laminar diffusivity
 ζ = fraction of energy generated in gas which is deposited in solids
 λ = ratio of jet injection radius to total radius of vortex tube
 μ = viscosity
 ν = kinematic viscosity
 ξ = ratio of radiated energy to that deposited in the gas
 ω = dimensionless criticality parameter defined in Eq. 21
 ρ = density
 Σ_f = macroscopic fission cross section
 σ_f = microscopic fission cross section
 ω = temperature exponent of diffusion coefficient

Subscripts

$*$ = turbulent value
 $'$ = normalized with respect to value at position of maximum w
 $(-)$ = normalized by division by r_0
 a = evaluated at outer edge of parabolic w profile
 c = evaluated at inner edge of parabolic w profile
 e = maximum exit radius of vortex
 i = injection conditions
 m = evaluated at position of maximum w
 0 = evaluated at outer radius of vortex
 s = standard conditions
 $1, 2$ = referring to light and heavy gas, respectively

References

- 1 Kerrebrock, J. L. and Meghreblian, R. V., "An Analysis of Vortex Tubes for Combined Gas-Phase Fission Heating and Separation of the Fissionable Material," Oak Ridge National Laboratory, ORNL CF-57-11-3, rev. 1, April 11, 1958, *declassified December 22, 1959*; see also: Kerrebrock, J. L. and Meghreblian, R. V., "Vortex Containment for the Gaseous Fission Rocket," to be published in *J. Aerospace Sciences*.
- 2 Keyes, J. J., "An Experimental Study of Gas Dynamics in High Velocity Vortex Flow," *Proc. Heat. Trans. and Fluid Mech. Inst.*, Stanford University Press, Stanford, Calif., June 1960.
- 3 Ragsdale, R. G., "NASA Research on the Hydrodynamics of the Gaseous Vortex Reactor," NASA TN D-288, Sept. 1960.
- 4 Keyes, J. J. and Dial, R. E., "An Experimental Study of Vortex Flow for Application to Gas-Phase Fission Heating," Oak Ridge National Laboratory, ORNL-2837, April 16, 1960, *Declassified*, Sept. 13, 1960.
- 5 Chapman, S. and Cowling, T. G., *The Mathematical Theory of Non-Uniform Gases*, Second Ed., Cambridge University Press, N. Y., 1958.
- 6 Einstein, H. A. and Li, H. L., "Steady Vortex Flow in a Real Fluid," *Proc. Heat Trans. and Fluid Mech. Inst.*, Stanford University Press, Stanford, Calif., 1951.
- 7 Rosenzweig, M. L., "Velocity Recovery and Shear Reduction in Jet-Driven Vortex Tubes," Aerospace Corp. Rep. TDR-94 (1203-01)TN-1, March 10, 1961.
- 8 Kerrebrock, J. L. and Lafyatis, P. J., "Analytical Study of Some Aspects of Vortex Tubes for Gas-Phase Fission Heating," Oak Ridge National Laboratory, ORNL CF-58-7-4, July 21, 1958, *Declassified*, Sept. 2, 1960.
- 9 Evans, R. D., *The Atomic Nucleus*, McGraw-Hill, 1955.
- 10 Kendall, J., Jet Propulsion Laboratory, private communication.
- 11 Rosenzweig, M. L., "The Vortex Matrix Approach to Gaseous Nuclear Propulsion," ARS preprint 1735-61, May 1961.

1961 ARS Meeting Schedule

Date	Meeting	Location	Abstract Deadline
Aug. 7-9	Guidance and Control Conference	Palo Alto, Calif.	Past
Aug. 16-18	International Hypersonics Conference	Cambridge, Mass.	Past
Aug. 23-25	Biennial Gas Dynamics Symposium	Evanston, Ill.	Past
Oct. 2-7	XIIth International Astronautical Congress	Washington, D.C.	Past
Oct. 9-15	ARS SPACE FLIGHT REPORT TO THE NATION	New York, N.Y.	Past

Send all abstracts to Meetings Manager, ARS, 500 Fifth Ave., New York 36, N.Y.

Propulsion System Using a Cavity Reactor and Magnetohydrodynamic Generator

RICHARD J. ROSA¹

Avco-Everett Research Laboratory
Everett, Mass.

This report outlines a general propulsion scheme using these two devices. In spite of being in a sense a "nuclear electric" system, accelerations well in excess of 1G seem possible. A specific system is described including estimates of its size, weight and performance. It appears that it might compete very well with other high thrust systems either chemical or nuclear for boosting vehicles from the ground into satellite orbit and beyond.

THE POSSIBILITIES for magnetohydrodynamic (MHD) power generation have been studied at the Avco-Everett Research Laboratory for some time, and the nuclear cavity reactor has received the attention of a number of people within the AEC and elsewhere.

The cavity reactor by itself does not make a good nuclear rocket nor does the MHD generator by itself make possible an ultra high performance electric power supply. In combination, however, they might achieve both of these results.

It appears that a cavity reactor (1,2,3)² might by itself make a very good rocket (4) except for the fact that it does not seem possible to have more than a very small fraction of the uranium undergo fission as it passes through the cavity. Hence, such a rocket would dump very large quantities of unused uranium overboard in the exhaust.

This objection would, of course, be dismissed if the uranium could be removed from the reactor's exhaust and recycled. However, at temperatures of interest for propulsion, filtering of uranium directly from the hot exhaust of the reactor with reasonable collection efficiency and enthalpy loss is clearly a very difficult, if not impossible, job. This is especially true if all or part of the uranium is in vapor form.

This report discusses an alternative way of using a cavity reactor. It is still in essence a filtering process since it removes the uranium from the exhaust but lets the propellant and the energy flow through. Although a relatively complex system, it might still be capable of a thrust many times its own weight.

Outline of the System

A block diagram of the system which is being proposed is shown in Fig. 1. The reactor exhaust is first cooled by doing work in an expansion engine of some sort such as a turbine or MHD generator. With sufficient cooling, efficient uranium collection could become possible. In particular, any uranium vapor present could condense and heat transfer losses in the collecting device reduced to acceptable levels. This collector

might be modeled after the cyclone separators and scrubbers presently used in industry to remove dust or liquid droplets from gases and vapors.

After the gas has been cleaned, it is fed into a heating, compressing or accelerating device which returns to it the energy that had previously been extracted.

It is highly unlikely that such a multiple energy conversion process could be carried out to advantage using conventional machinery such as turbogenerators and compressors. The temperatures of interest in this application are much too high for these devices and, in any case, their weight would be

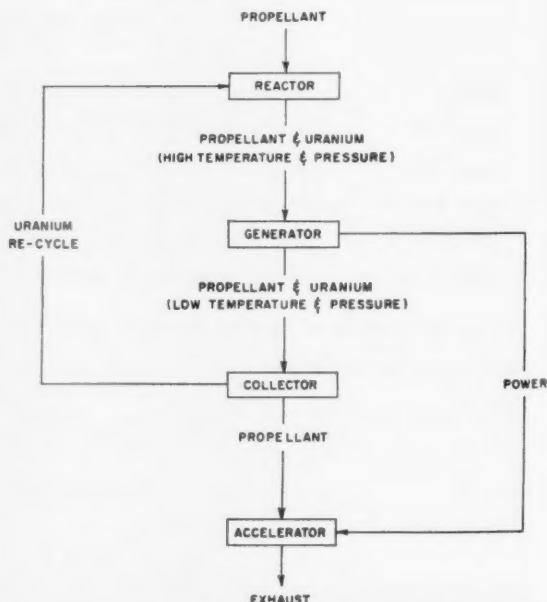


Fig. 1 Open cycle propulsion system using an MHD generator and nuclear cavity reactor

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ Principal Research Scientist. Member ARS.

² Numbers in parentheses indicate References at end of paper.

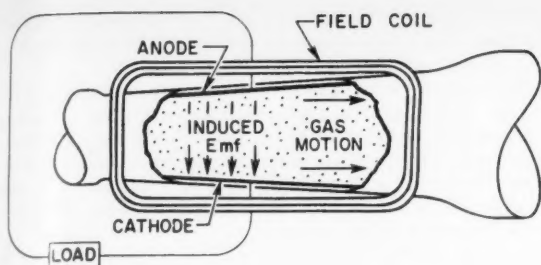


Fig. 2 D.C. magnetohydrodynamic generator

so great as to make the whole system quite unattractive, or suitable only for applications where low thrust-to-weight ratios are tolerable.

However, MHD generators and accelerators (5) should be able to withstand high temperatures and may produce very large amounts of power per pound of weight. These potential characteristics stem from the relative absence of solid material parts, either stationary or moving, in the basic energy conversion process.

Generator Performance

A sketch of a simple type of MHD generator is shown in Fig. 2. It consists of a duct through which the gas flows, coils which produce a magnetic field across the duct and electrodes at the top and bottom of the duct. These electrodes serve much the same purpose as the brushes in a conventional generator. The gas, by virtue of its motion through the magnetic field, has an emf generated in it which drives a current through it, the electrodes and the external load. As is the case in a turbine, this gas motion occurs because of a pressure gradient through the device. The total pressure drop then is a function of the amount of work which the gas does and the extent to which it does it in an isentropic or reversible manner.

It is, of course, necessary that the gas be an electrical conductor, that is, be partially ionized. It appears that at the temperatures required for this application the gas will become sufficiently ionized if seeded with a one per cent mole fraction more or less of potassium. Calculated values of the resulting conductivity are shown in Fig. 3.

An experimental MHD generator was constructed and successfully operated at the Avco-Everett Research Laboratory nearly two years ago. This device, which is shown in Fig. 4, produced somewhat over 10 kw of electric power and behaved largely in accordance with theoretical expectations (6). MHD generators are now being actively considered by Avco and by others for large scale commercial power generation from either nuclear or fossil fuels (7).

Fig. 5 shows estimates, of the specific power output² of an MHD generator as a function of the electrical conductivity of the gaseous working fluid. Temperatures corresponding to the conductivity values are also shown which are more or less typical for a gas seeded with potassium vapor. Lines of constant rated power output are shown. These lines are based on a mean field strength of 10,000 gauss, a mean gas velocity of 1000 m per sec, and a field coil excitation power equal to 1% of the rated power output. Fig. 6 shows the factors by which the specific output changes as these parameters are varied.

For purposes of orientation, the overall system specific output required for 1G acceleration with an impulse of 1000 sec

² Output per unit field coil mass is shown on the graph. However, unless very high gas pressures are used, the field coil should be the dominant weight.

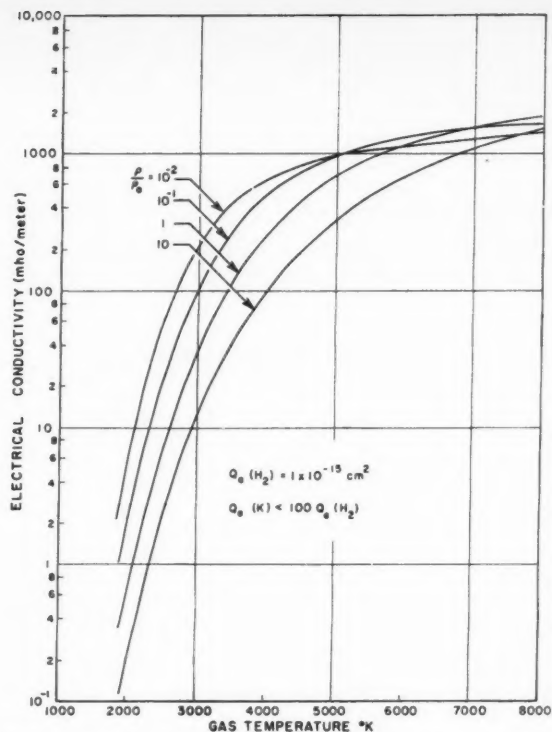


Fig. 3 Electrical conductivity of hydrogen plus 1% potassium

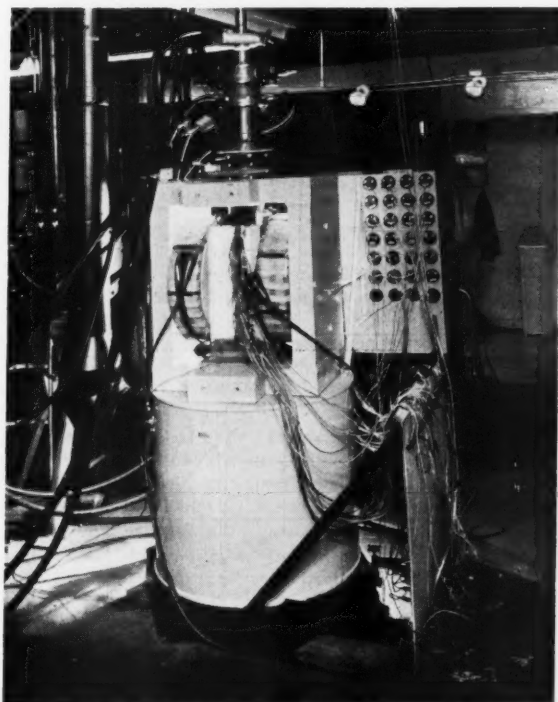


Fig. 4 Experimental 10 kw MHD generator facility

and the output obtainable from a turbogenerator are marked on the ordinate scale. It can be seen that for sufficiently high gas conductivities, high powered MHD devices should be capable of specific outputs substantially in excess of both of these points. With this as a possibility, a multistep filtering and energy conversion process as described above becomes compatible with high thrust.

System Characteristics

In order to give some idea of what the size, weight, losses, etc. in a scheme like this might be, the rough outlines of a

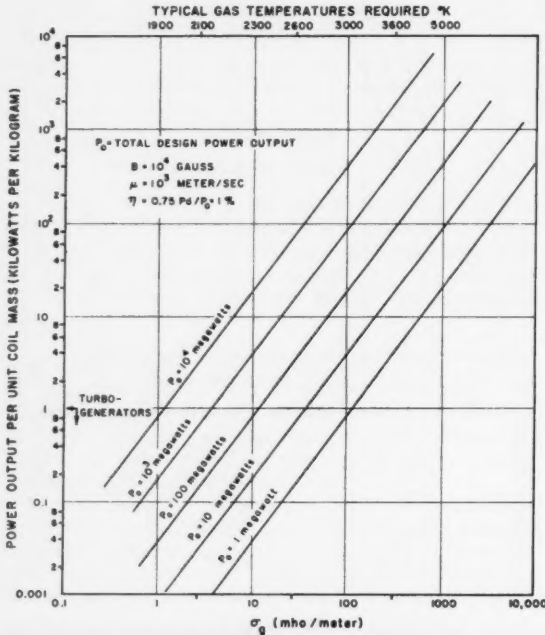


Fig. 5 Characteristics of an MHD generator

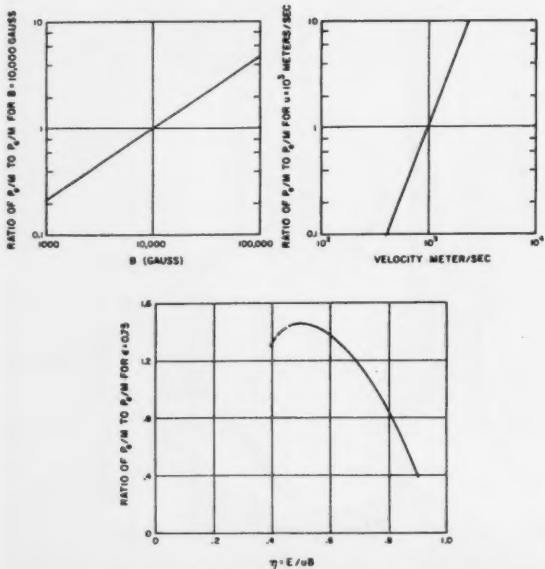


Fig. 6 Variation of P_0/M for changes in the parameters used in Fig. 5

specific system will be sketched out. In an appendix some of the basic design considerations and ground rules for such a system are discussed.

The characteristics of both cavity reactors and MHD generators dictate that a high performance system be large. The smallest vehicle for which this scheme is appropriate would probably have an initial weight of several hundred tons. However, it is probable that any nuclear propelled vehicle that is manned and hence carries shielding will have to be of about this size anyway.

A configuration of a system is sketched in Fig. 7. It is physically quite large; however, much of it consists of thin walled structures containing gas at relatively low pressure and temperature. This unit is intended to produce about 300 tons of thrust with a specific impulse somewhat greater than 1000 sec. Hence, for example, it should be capable of putting about 100 tons into a close satellite orbit. With some increase in complexity and weight, impulses approaching 2000 sec might be possible with a system of this type.

Reactor

It has been assumed that the reactor operates at a pressure of 100 atm and produces a temperature of 3800 K in a working fluid consisting of hydrogen plus 1% uranium and 1% potassium vapor. It is estimated that for criticality such a reactor needs to be about 2 m in diameter and surrounded by a blanket of D_2O weighing about 7 tons.

Generator

A simple linear MHD generator configuration has been assumed as shown in Fig. 2. The gas expands through a pressure ratio of 100 in the generator (see Mollier diagram, Fig. 8) coming out at 2000 K and one atmosphere pressure. The values of electrical conductivity that this exhaust gas will have can be calculated with reasonable confidence and are shown in Fig. 3. It is then possible to work out the dimensions of the generator and the magnetic field strengths which

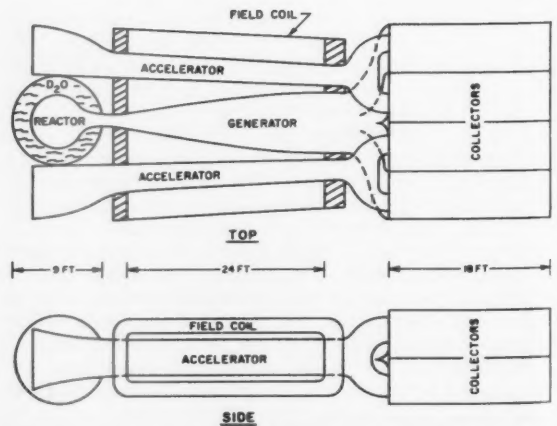


Fig. 7 Sketch of a 300 ton thrust nuclear MHD open cycle propulsion system

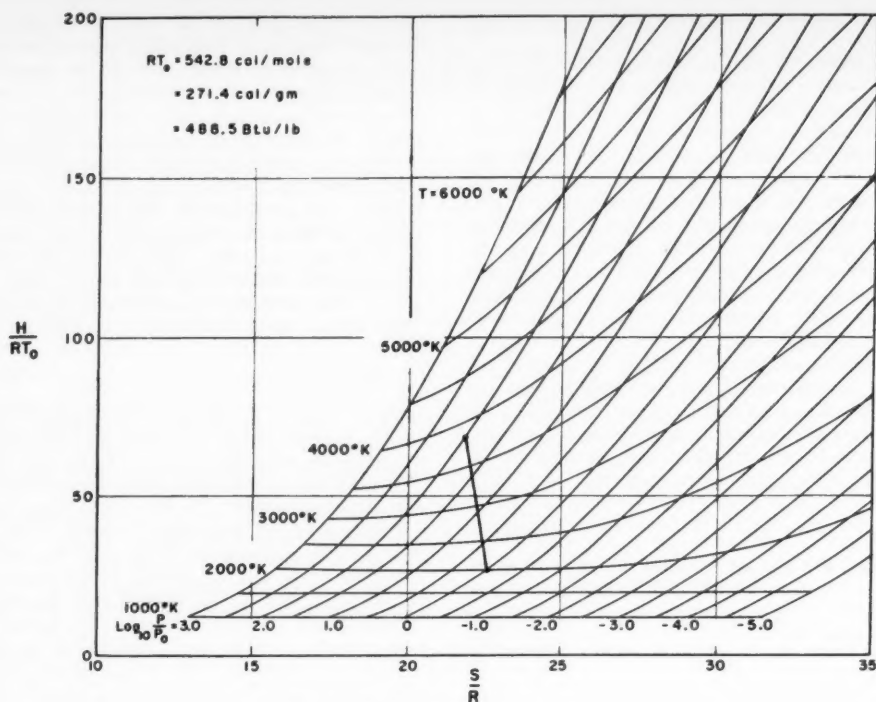


Fig. 8 Hydrogen Mollier diagram

should be employed. From this, the weight, dimensions and power consumption of the coil to produce the field can be estimated.

The duct of this generator has a square cross section and is 8 m long. It is designed for a gas flow of 1.4×10^5 moles per sec and a power output of 14,000 megawatts. This power, although enormous compared to present rates of electric power generation, is not so very much greater than the thrust power of present chemical rockets. There does not seem to be any fundamental engineering consideration that would prohibit an MHD generator, as described, from converting gas enthalpy to electric power on this or an even larger scale.

If 300 K values for copper conductivity are used, the coil weight for this generator is estimated at 12 tons and its power consumption at 110 megawatts. However, there is more than enough cooling capacity in the liquid hydrogen as it comes from the propellant tanks to keep the coils below about 70 K and gain thereby a factor of at least 10 in copper conductivity. The coil might then weigh about 4 tons and consume about 40 megawatts.

Collectors

At the 2000 K generator exit temperature, the vapor pressure of uranium is 10^{-6} atm (or less if a refractory compound is used). Hence, any vapor should have condensed into droplets. In Fig. 7, eight cyclone collectors are shown operating in parallel, each collector being about 2 m in diameter, 6 m long, and weighing about half a ton. It is estimated that such a collector should remove essentially all particles whose diameter is greater than about 10μ . If necessary it might

be feasible to add an additional scrubbing process to get out smaller particles.

As it leaves the collector the propellant should contain ideally no more than a 10^{-6} mole fraction or about 10^{-4} by weight of uranium. Putting about 100 tons into a satellite orbit would then result in a loss of about 40 lb of uranium. This would seem to be well within permissible limits.

Accelerators

The propellant after leaving the collectors passes into the accelerator. Several types of accelerators might be used. However, a magnetohydrodynamic device would be preferable to a simple plasma jet, fundamentally because with the former, one should be able to obtain a more nearly isentropic acceleration process, losing less energy in dissociation of the hydrogen and maintaining a higher exit pressure.

The device sketched in Fig. 7 is intended to accelerate the gas more or less isothermally up to a Mach number of 3 at about 2000 K.⁴ See (8,9). This is consistent with the amount of electric power that the gas produced in going through the generator and results in a specific impulse of about 1000 sec. If the ambient pressure is low enough the gas may then be expanded in a nozzle to a higher Mach number with a possible increase in specific impulse up to about 1200 sec.

In the accelerator shown, the field coil weighs 2 tons, and requires 30 megawatts of electric power. This again assumes values of copper conductivity about one-tenth the 300 K value.

⁴ It is assumed that the gas still contains about 1% equivalent potassium vapor to give it the necessary conductivity.

Cooling

The cooling requirements for this system are met by using the propellant as coolant before and as it enters the reactor. This cooling capacity seems to be adequate. Of the order of 10^8 cal/sec will be dissipated in the D_2O due to moderation of neutrons. The field coils of the generator and accelerator dissipate a total of 70 megawatts. Wall heat transfer in the generator is estimated to be a total of 450 megawatts, and in the accelerator, about 100 megawatts. It would be desirable and might be possible to leave the cyclone collector walls uncooled except by their own radiation and even to inhibit this radiation so that they approached the 2000 K gas temperature. If, however, cooling were required, an additional heat loss of about 400 megawatts might result.

The total cooling requirement for coils and walls (excluding the reactor and generator entrance walls where internal film cooling is probably necessary) is thus about 1500 megawatts or 3.4×10^8 cal per sec. Since the hydrogen flows at the rate of 1.4×10^6 moles per sec, it will, in satisfying the cooling requirement, be heated from 20 K to about 400 K.

Summary

As stated earlier this system produces about 300 tons of thrust at a specific impulse of 1000 and should thus be able to put something like a 100-ton vehicle into a close satellite orbit. Of this 100 tons the propulsion system would constitute the following:

7 tons	reactor
4 tons	generator
4 tons	collectors
2 tons	accelerator
17 tons ⁶	total propulsion system

In addition, if personnel as well as freight are carried, a

⁶ If it is in fact not practical to maintain the field coils at a temperature below about 70 deg Kelvin, the weight of the propulsion system might grow to about 30 tons.

considerable allowance for shielding must, of course, be included. Perhaps 20 tons would suffice if the propellant tanks formed part of the shield. This leaves about 80 tons (60 tons less shielding) for structure and payload. The power supply itself would constitute useful payload for some types of missions. In particular, it could, after attaining an orbit, be converted to a closed cycle system for powering an ultra high impulse accelerator, such as an ion rocket, on interplanetary flights. Its specific power output would be considerably reduced by this conversion, largely because of the necessity for a radiating heat sink. However, partly because high (although materials limited) sink temperatures could be used, the specific output should be much higher than attainable with more conventional generating equipment.

Conclusions

A propulsion system like the one outlined above would make the placing of very large payloads into satellite orbit much easier than it is ever apt to be with chemical rockets. It would thus make extensive, manned ventures into space much more feasible. However, quite apart from the merits or deficiencies of this particular scheme is the example it provides of the potentialities of its two major components. The gaseous cavity reactor and the MHD generator are both devices in which temperature and power density are not fundamentally limited by the properties of solid materials. In combination they seem to make possible the generation of electric power on a scale hitherto undreamed of. It is hard to doubt that the realization of such a possibility would have great engineering significance.

Appendix

General Design Considerations

There are a number of basic quantities and relations that govern the design and performance of the system described. Some of these are:

- 1 The vapor pressure of uranium (or of its compounds if

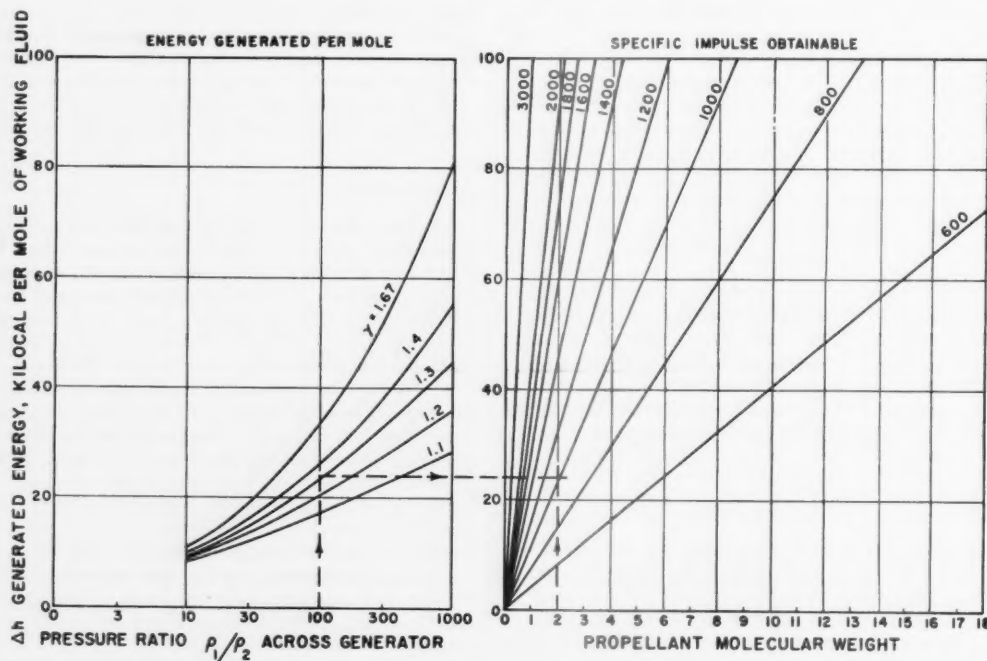


Fig. 9 Energy generation and attainable specific impulse for a generator efficiency of 80% and exit temperature of 2000 K

suitable ones exist) and its latent heat of vaporization.

2 The electrical conductivity that can be obtained in gases as a function of temperature (the temperature at which uranium or suitable compounds have a low vapor pressure being of particular interest).

3 The uranium density necessary for criticality in reactors of reasonable size and weight.

4 The pressures and enthalpies permitted by structural and heat transfer considerations.

5 The desired thrust density and/or operational ambient pressure or altitude limits.

The specific impulse which the system develops will be determined by the molecular weight of the propellant and the energy delivered by the generator for each mole of propellant-vapor mixture that passes through it. This energy is given by

$$\Delta h = \int_{T_1}^{T_2} C_p dT \approx \bar{C}_p (T_1 - T_2) \quad [1]$$

where \bar{C}_p is an effective heat capacity averaged over the temperature range $(T_1 - T_2)$. As in the case of a turbine, the gas expands approximately isentropically as it passes through an MHD generator. Hence

$$\frac{T_1}{T_2} = \left(\frac{P_1}{P_2} \right)^{(\gamma-1)\epsilon/\gamma} \quad [2]$$

where gamma is the specific heat ratio and epsilon is a number less than one that defines the "efficiency" or entropy generation in the device. (Approximately, $\epsilon \approx$ load impedance \div generator plus load impedance.) Thus in terms of pressure

$$\Delta h = \frac{\gamma}{\gamma-1} \gamma R T_2 \left[\left(\frac{P_1}{P_2} \right)^{(\gamma-1)\epsilon/\gamma} - 1 \right] \quad [3]$$

A bottom temperature T_2 must be chosen that will result in a sufficiently low vapor pressure for uranium and a sufficiently high electrical conductivity for the gas. P_1 must be high enough for reactor criticality and evidently should be higher if structures and heat transfer will allow. P_2 should evidently be as low as collection efficiency and thrust density requirements will allow.

Uranium Collection Efficiency and Gas Conductivity

If a rocket has a specific impulse of about 1000 sec and its exhaust contains one atom of uranium per million of propellant (or about one part in ten thousand by weight), it will, in accelerating to the satellite velocity (26,000 ft/sec), lose about one pound of uranium per ton of final weight. This would seem to be a very acceptable loss rate and possible of attainment if separators for this application can be made to perform as well as those presently used in industry. If we desire operation at or near sea level, then the total gas pressure at the generator exit must be in the vicinity of one atmosphere. Ref. (10) gives vapor pressure vs. temperature for uranium. It indicates that for purification of the vapor to one part in a million the temperature at the generator exit must be about 2000 K or less. If a refractory uranium compound is used higher temperatures might be permissible. However, enthalpy loss and structural problems in the collectors will increase with the temperature. Present indications are that an exit temperature much less than 2000 K is not possible since it appears that with any combination of gas and seed material the electrical conductivity will quickly drop to unusable values as one goes to lower temperatures.

Starting Pressure—Reactor Criticality

Estimates based on information (1,2,3) indicate that a U_{235} particle concentration of 3×10^{18} per cubic centimeter is just about the lowest concentration that results in a reasonable size and weight for the reactor. Since gas temperatures of the order of 10 times room temperature would be asked for, this density translates into a required uranium partial pressure in the reactor of the order of one atmosphere.

In determining what fraction of the working fluid in the reactor is uranium, it must be kept in mind that uranium has a high latent heat of vaporization (110 kcal/mole). If uranium in the vapor phase constitutes much more than about 1% of the working fluid it will, as it condenses, appreciably lower the specific heat ratio of the gas. This will make the energy extraction process more difficult as can be shown from Eq. 3.

The above two considerations indicate that a total gas pressure of the order of one hundred atmosphere is desirable in the reactor. Even higher pressures are, of course, desirable from the point of view of getting the maximum specific impulse. However, with higher pressures one could expect heat transfer problems which would be considerably more severe than those in present day chemical rockets.

Starting Pressure—Attainable Specific Impulse

Fig. 9 shows the power per mole that may be generated as a function of pressure ratio for various values of gamma. It has been assumed that the temperature at the end of expansion is 2000 K and the efficiency factor " ϵ " is equal to 0.8. Also shown is the specific impulse attainable as a function of the energy generated and the propellant molecular weight. The latter graph assumes 100% utilization of the generated power. However, it does not include utilization of the enthalpy which the gas still contains (15 kcal/mole) at 2000 K. Since it should be possible to convert at least some of this energy into directed motion the graph is probably a fairly reasonable estimate of attainable specific impulses.

Impulses appreciably higher than shown here might be obtained if the 2000 K generator exit temperature could be increased. It might be possible, for example, to regeneratively cool the exit gases with the propellant as it comes from the tanks. The generator exit temperature could then be 2000 K plus whatever ΔT one could afford in the regenerator. For example, in the system outlined above, after cooling the field coils, duct walls, and the moderator of the reactor, the propellant temperature was only about 400 K. It thus might have been used to regeneratively cool the exit gases through a ΔT of perhaps 1000 K making possible a generator exit temperature of 3000 K. This would result in about a 20% increase in specific impulse.

References

- 1 Bell, L. A., Los Alamos Rep. L.A. 1874, 1955.
- 2 Safanov, G., Rand Rep. R.M. 1835, 1955.
- 3 Colgate, E. A. and Aamodt, R. L., *Nucleonics*, vol. 15, Aug. 1957.
- 4 Bussard, R. W., *JET PROPULSION*: vol. 28, April 1958, Shepherd and Cleaver, *J. Brit. Interplanetary Soc.*, vol. 7, no. 5, p. 185; no. 6, p. 234, 1948.
- 5 Rosa, R. J. and Kantrowitz, A. R., "Magnetohydrodynamic Energy Conversion Technique," proceedings of Seminar on Advanced Energy Sources and Conversion Techniques, Pasadena, Calif., Nov. 1958.
- 6 Rosa, R. J., "Physical Principles of Magnetohydrodynamic Power Generation," *Physics of Fluids*, vol. 4, Feb. 1961, pp. 182-194.
- 7 Sporn, P. and Kantrowitz, A., "Large Scale Generation of Electric Power Application of the Magnetohydrodynamic Concept," *Power*, Nov. 1959.
- 8 Wood, G. and Carter, A. F., "Dynamics of Conducting Gases," proceedings of the Third Biennial Gas Dynamics Symposium, Northwestern University Press, 1960, pp. 201-212.
- 9 Rosa, R. J., "Engineering Magnetohydrodynamic and Shock Wave Spectroscopy," thesis submitted in partial fulfillment of the requirement for the degree of Ph.D., Cornell University, 1956.
- 10 Brewer, L., "The Thermodynamic and Physical Properties of the Elements," *The Chemistry and Metallurgy of Miscellaneous Materials*, by L. L. Quill, McGraw-Hill, N. Y., 1950, pp. 13-19.

Photochemistry and Space Power Generation

J. N. PITTS Jr.¹

University of California
Riverside, Calif.

J. DAVID MARGERUM²
and Wm. E. McKEE³

Sundstrand Corporation
Pacoima, Calif.

Some of the fundamental aspects of photochemistry are reviewed with reference to the direct use of solar radiation for the initiation of chemical reactions from which power can be generated. Photoregenerative fuel cells and photogalvanic cells are defined, and they are discussed in terms of criteria of usefulness, calculations of theoretical efficiencies and limiting factors. The nitrosyl chloride system is used as an example to illustrate the theoretical calculations of efficiency and maximum power available from a solar regenerative fuel cell. The development of a fuel cell operating on nitric oxide and chlorine is reported. The competitive process of nitric oxide-chlorine recombination following the photodecomposition of nitrosyl chloride is also considered. The problem of direct energy storage in photogalvanic batteries is discussed, and some areas of research for the development of photogalvanic cells are recommended.

IN THE decade following World War II there gradually arose a widespread interest in the development of practical systems for terrestrial solar energy conversion. Provision of power for remote portions of Earth's surface was recognized as a significant existing problem. Development of solar conversion devices to join nuclear power plants in replacing Earth's dwindling supply of fossil fuels was considered vitally important but more as a long term matter.

With the launching of Sputnik I on October 4, 1957 efficient solar energy conversion became recognized as a matter of immediate importance in international political as well as scientific circles. The usefulness of artificial satellites obviously depends to a great extent upon the lifetime of the systems providing electrical power to the telemetering equipment. The solar radiation band is at present the only type of energy flux in space sufficiently intense to be a worthwhile source of natural energy of indefinite duration.

Research on methods of efficiently converting the 1400 w/m² of primary solar energy in space has a high priority in our space vehicle program. To date, only the photovoltaic silicon junction cell has been successfully employed in satellites, performing very well indeed in Vanguard I and the Pioneer V. Such photovoltaic cells have at present a practical efficiency of about 10% of the solar energy. These "solar batteries" are expensive and subject to certain theoretical and mechanical limitations. Thus, concurrently there is widespread research aimed at exploring other principles of energy conversion. These include thermionic devices, thermoelectric units, solar heat engines, photogalvanic solar cells and solar regenerative fuel cells. This paper deals with fundamental photochemical principles and processes relative to the latter two types of power generation for space vehicles. These are the farthest back in the development of devices, yet they are potentially of great value for both terrestrial and space applications. We shall discuss the photochemical theory of solar regenerative fuel cells, particularly one employing nitrosyl chloride (NOCl), and shall

consider some future areas of research in the photochemistry of photogalvanic systems.

For a detailed consideration of the entire field of photochemical processes related to the utilization of solar energy the reader is referred to Calvert's discussion on this subject, (1).⁴ A concise review has also been published by Rosenberg (2), and photogalvanic cells have been reviewed by Sancier (3).

Solar regenerative fuel cells are recyclable systems in which the "working fluid" is irradiated by sunlight and photodissociated into products which can be separated and stored. Later they are allowed to react in a suitable fuel cell to produce electrical power and the original working fluid. The latter can be rephotolyzed to initiate another power cycle. The potential storage capability of this type of solar cell is one of its most attractive features.

A photogalvanic cell can be defined as a system containing electrodes and electrolyte in which irradiation of the system causes a reversible chemical change that produces an electrical potential. The photolysis products do not need to be separated, as contrasted with a photoregenerative fuel cell. In current usage photogalvanic systems are distinguished from photovoltaic systems in that the latter do not contain an electrolyte and are solid state devices. (Photovoltaic cells are not considered in this paper.) The most widely studied type of photogalvanic cell contains two electrodes and an aqueous solution containing an organic dye and inorganic ions of a suitable redox couple. Illumination of the solution surrounding one electrode causes photooxidation reduction reactions which set up an E.M.F. between this electrode and the electrodes in the dark.

In both types of systems, it is possible theoretically to recycle them indefinitely, the electrical power arising from the conversion of part of the solar energy into chemical energy. The prime prerequisite for a recyclable system is that the overall photochemical process be "endothermic," i.e., that the chemical energy of the products be greater than that of the reactants. Such a restriction greatly limits the number of feasible photolysis systems. Suitable "uphill" photochemical reactions are scarce. Some reactions for possible

Presented at the ARS Semi-Annual Meeting, Los Angeles, Calif., May 9-12, 1960.

¹ Professor of Chemistry.

² Research Specialist.

³ Chemical Section Manager.

⁴ Numbers in parentheses indicate References at end of paper.

use in future large-scale terrestrial plants have been evaluated, but space applications present significantly different aspects.⁵ These include the following:

1 The system should be operable under conditions of zero gravity.

2 The cyclic system should be essentially 100% reversible, with no constituent being used up or requiring replacement.

3 An "endothermic" photochemical process for a recyclable fuel cell or photogalvanic battery does not have to be economically competitive with other energy sources, as is the case for the proposed large scale commercial conversion of solar energy. Thus, reactions which must be eliminated from consideration because they involve chemicals and apparatus too expensive for a large scale project may be entirely suitable for space applications. It is energy vs. pounds of weight that becomes of utmost importance.

4 Absorption of most of the ultraviolet solar radiation below 3100 Å by Earth's atmosphere limits the effective region for large scale commercial photochemical processes to the range from about 3100 to 8000 Å. Such a restriction is absent in the case of a photolysis system for use in a satellite in orbit, and the probability of finding suitable photoreactions is increased since solar radiation in the region below 3100 Å is available (although it is only a small fraction of the total radiant energy from the sun, see Table 1).

Despite these differences, the published research on possible terrestrial processes for conversion of solar energy to chemical energy is clearly a useful source of information on photoreactions of potential interest in regenerative fuel cells and photogalvanic cells for space applications. Calvert's tables (1) of the theoretical efficiency of endothermic photochemical reactions is a particularly useful reference source.

Theoretical Aspects of the Problem

Certain fundamental considerations arise when dealing with the conversion of radiant energy to chemical energy in an endothermic photochemical reaction. These can be summarized as follows:

Absorption of Radiation

In order for radiation to be effective in producing a chemical change, it must be absorbed either by the molecule to be dissociated, or by a "sensitizer" added to the solution which absorbs the light at shorter (and sometimes longer) wave lengths and then transfers energy to the molecule to be decomposed. When a photon is absorbed, the energy of the absorbing molecule is increased by the amount $h\nu$ or hc/λ where ν is the frequency in cm^{-1} , λ is the wave length in cm, h is Planck's constant (6.62×10^{-27} erg sec) and c is the velocity of light (3×10^{10} cm/sec).

Photochemists find it more convenient to think in terms of an Avogadro of light quanta, i.e., 6.02×10^{23} photons. This is called an einstein of radiation and is written $Nh\nu$. A useful conversion factor giving the energy of mole of quanta is

$$Nh\nu = \frac{2.86 \times 10^5}{\lambda(\text{\AA})} \text{ kcal} = \frac{11.35}{\lambda(\text{\AA})} \text{ Btu}$$

Since Einstein's law of photochemistry states that simultaneous absorption by a molecule of more than one photon of light is improbable, the value of $Nh\nu$ for a given wave length represents the amount of radiant energy taken up by the absorbing molecules.

Absorption of radiant energy does not necessarily lead to overall lasting chemical changes. Fluorescence, deactivation, back reactions of the products, etc. also may occur and these lower the photochemical efficiency of the process.

Usually only light in the region 1500 to 8000 Å is of photo-

⁵ A recent paper by Zwick and Zimmerman⁽⁴⁾ reviews space vehicle power systems.

Table 1 [After Friedman (5)] Per cent of incident energy below a given wave length

Wave length Å	Per cent incident energy below wave length
2,500	0.2
3,000	1.2
4,000	9.0
5,000	24.0
6,000	37.0
7,000	49.0
8,000	58.0
9,000	65.0
10,000	71.0
12,000	80.0
15,000	88.0
20,000	94.0

chemical interest. This corresponds to transitions between molecular electronic states. At wave lengths below 1500 Å ionization often occurs. Beyond 8000 Å the absorbed photons usually cause an increase in the vibrational and rotational energy of the molecule; the absorbed energy is too weak to break chemical bonds and the ultimate effect is simply to heat the system. About 58% of the solar radiation falls in the photochemically useful region below 8000 Å, while the rest is not utilizable by known photochemical processes. Thus, the photochemist is concerned with the distribution of energy as a function of wave length. Table 1 shows the distribution above Earth's atmosphere.

This energy distribution is approximately that of a black body at a temperature of about 6000 K, with the maximum spectral intensity at approximately 5000 Å (6). The total radiant flux in space at the edge of Earth's atmosphere is about 2.0 cal/cm²/min, corresponding to 1.4 kw/m² or 1.5 hp/yd². (This flux, of course, varies with the distance from the sun, and has values of 2.6 and 0.6 kw/m² at Venus and Mars respectively). At the Earth's surface, losses due to scattering, absorption, etc. reduce the average flux reaching the surface to about 1.0 kw/m² on clear days or approximately 0.7 kw/m² for a long time average.

Criteria for Useful Endothermic Photochemical Processes

The requirements for potentially useful photochemical processes can be summarized as follows:

1 Strong absorption of light by the reactant over a wide range in order to absorb an appreciable fraction of the solar energy. This is also necessary if the thickness of the absorption cell is to be kept reasonably short.

2 Weak absorption of light by the products. This is necessary or the "inner filter" effect will occur and much of the incident radiation will be absorbed by the products and, in effect, wasted.

3 There should be a high quantum yield Φ for the photo-decomposition.

$$\Phi = \frac{\text{No. of molecules of product}}{\text{No. of quanta absorbed}}$$

$$= \frac{\text{No. of molecules of product}}{I \times Q_m}$$

where I = incident intensity in quanta/sec; Q_m = fraction of light absorbed. Generally an endothermic reaction has a maximum quantum yield of unity, except for special cases e.g., nitrosyl chloride has a maximum quantum yield of two for the formation of NO.

Concurrently with the quantum yield, the concept of the mechanism of a photochemical process must be considered. Most photochemical reactions occur in several steps called elementary processes. The absorption of light followed by the first chemical act is called the *primary process*. Einstein pointed out that the primary quantum yield of decomposition ϕ could not exceed unity. However, if the products of the primary process are reactive, *secondary reactions* which are usually thermal in nature may occur, giving an overall quantum yield well in excess of unity. As an example, the broad aspects of the mechanism of the photolysis of gaseous mixtures of hydrogen and chlorine are given below:

Primary process

Chain initiating.

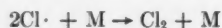
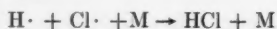


Secondary reactions

Chain propagating.



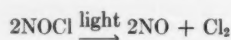
Chain terminating.



This is an example of a chain reaction of explosive character in which ϕ HCl may exceed 10^5 – 10^6 (10^5 – 10^6 molecules of HCl are formed per quantum of light absorbed by the chlorine molecule). The symbol *M* stands for a "third body" which must be present to siphon off the excess vibrational energy of the HCl formed in the fourth procedure and Cl_2 in the last.

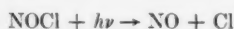
Such chain reactions are of great importance theoretically and commercially, but of necessity the overall processes must be exothermic. Hence, they cannot be used for recyclable solar conversion devices. On the other hand, the photolysis of nitrosyl chloride is an example of an endothermic process which has attractive features for use in a solar regenerative fuel cell. It absorbs strongly in the region from 6300 to 3500 Å, and as shown by Kistiakowsky it photodecomposes with an overall quantum yield of two for nitric oxide (NO) over the entire region (7).

The overall process can be written as

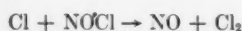


The detailed mechanism is more complex, but there is good evidence that it is a two step, free radical process (7,8,9,10).

Primary process



Secondary reaction



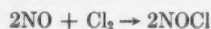
4 There should be a large increase in free energy ΔF for the overall reaction (the usual thermal exothermic reactions lead to a free energy decrease). The free energy is related to the E.M.F. of the chemical cell in which the photolysis products are combined by the equation

$$\Delta F = -\eta \mathfrak{F}V$$

where n = no. of Faradays, V = E.M.F. of the cell and \mathfrak{F} = 96,540 coulombs.

5 The reverse reaction of the products to form the reactants should be slow (or the products should be capable of being physically separated rapidly). Thus a critical factor

in the feasibility of NOCl as a fuel for a regenerative fuel cell is the rapid rate of the back reaction



particularly in the pure liquid. If this back reaction occurs at a place other than at the electrodes of the fuel cell, the light energy is wasted and appears only as heat.

Efficiency of an Endothermic Photochemical Process

Calvert (1) has defined a useful, if approximate, means of expressing the efficiency of an endothermic photochemical reaction. The quantity Q is a measure of the per cent efficiency of conversion of light energy to chemical energy at a given wavelength.

$$Q = \frac{(100)(\Delta F_{298}^\circ)(\Phi_A \cdot)}{E_\lambda}$$

where

ΔF° = standard free energy for the overall reaction in terms of kcal/mole of product *A*

Φ_A = quantum yield of product *A*

E_λ = wavelength of light in kcal/mole of quanta

Thus, in effect, Q represents the percentage of the radiant energy absorbed at a given wavelength which is available as chemical free energy. For example, the efficiency of NOCl photodecomposition at its maximum wavelength of visible absorption (about 6500 Å) can be calculated using a quantum yield of 2 and a ΔF of 4.9 kcal/mole as

$$Q_{6500\text{Å}} = \frac{(100)(4.9)(2)}{(2.86 \times 10^5/6500)} = 22.3$$

However, at the ultraviolet edge of the solar spectrum (about 3000 Å), photolysis of nitrosyl chloride has a much lower efficiency because the quanta have far more energy than that required for dissociation: $Q_{3000\text{Å}} = 10.3$.

The maximum electrical energy W_{max} theoretically available from a photochemical-electrochemical system is given by the following equation

$$W_{\text{max}} = \Phi I_a t N^{-1} n \mathfrak{F} V^\circ$$

$$W_{\text{max}} = 1.60 \times 10^{-10} \Phi I_a t n V^\circ \text{ w-sec/m}^2$$

where Φ is the quantum yield of the product reaction, $I_a \cdot t$ is the amount of light absorbed in quanta/m², N is Avogadro's number (6.023×10^{23} molecules/mole), n is the number of equivalents per mole for the electrochemical reaction, \mathfrak{F} is Faraday's number (96,540 coulombs/equivalent) and V° is the open circuit voltage of the cell.

In the limit it is theoretically possible to utilize all of the W_{max} as the output power for such solar energy conversion systems. However, in practice, it is impossible to deliver all of the electrical energy to an external load, so that a voltage of less than V° would actually be used, depending upon the cell characteristics and the voltage desired. Thus, it is also helpful to calculate the power theoretically available when the electrochemical cell is operated at its condition of maximum power output $P_{\text{max (cell)}}$. (It should be noted that this is not necessarily the best operated condition for the system as a whole.)

$$P_{\text{max (cell)}} = \phi I_a N^{-1} n \mathfrak{F} V_{1/2}^\circ$$

$$P_{\text{max (cell)}} = 1.60 \times 10^{-10} \phi I_a n V_{1/2}^\circ \text{ w/m}^2$$

where, I_a is the intensity of the light absorbed in quanta/sec m² and $V_{1/2}^\circ$ is one-half the theoretical open circuit voltage. (For any cell the maximum theoretical power output occurs when the internal resistance equals the external load, which corresponds to an operating voltage drop across the load of

one-half the open circuit voltage.) Again using NOCl as an example, the maximum electrical energy and power hypothetically available from the system may be calculated, using the following idealizing assumptions:

1 All of the incident light between 6300 and 2150 Å is absorbed by the NOCl, none by the NO or Cl₂.

2 The quantum yield of NO formation is two over the entire range.

3 No back reaction to reform NOCl occurs, except that which yields electrical power.

4 The voltage developed by the NOCl fuel cell is 0.21 v. (This is the theoretical voltage, and voltages of this value have been obtained.)⁶

Then since $I_a = 1.41 \times 10^{21}$ quanta/sec·m² for the 2150–6300 Å region, while $\Phi = 2$, $n = 1$ and $V^0 = 0.21$ v, the maximum chemical energy of the NOCl system theoretically convertible into electrical energy is

$$W_{\max} = (1.60 \times 10^{-19}) (2) (1.41 \times 10^{21}) (0.21) = 95 \text{ w-sec/m}^2$$

The theoretical efficiency of a photochemical-electrical system can then be defined in terms of efficiency of maximum energy convertible E_w

$$E_w = W_{\max}/W_s$$

where W_s is the total energy of the incident solar radiation. Thus, for nitrosyl chloride, the maximum theoretical efficiency E_w for utilizing the solar energy in the 2150–6300 Å range (about 562 w/m²) is 16.9%. The rest of the energy appears mainly as heat. The maximum theoretical efficiency for NOCl drops to 6.8% over the entire solar spectrum, about 1400 w/m², the decrease being due to its lack of absorption or lack of photodecomposition at wave lengths longer than about 6500 Å.

Sensitized photolysis can sometimes be used to increase the efficiency of a system by extending the wavelength range over which solar energy is absorbed. This is particularly important in such cases as the photodecomposition of water by the ultraviolet absorption of cerium and iron solutions (11). We suggest that it may be possible to apply photosensitization to increase the theoretical efficiency of the NOCl photolysis system. While the lack of absorption above 6500 Å is not a critical factor in this system, it should be noted that wavelengths up to 7518 Å provide enough energy to break the Cl–NO bond (which has a bond strength of only 38 kcal/mole). If sensitizers (such as nitrosyl metal chlorides) cause photolysis of NOCl to occur up to 7500 Å, then the maximum theoretical efficiency would increase from 7 to 10% of the solar energy.

Factors Tending to Reduce the Efficiency

A number of factors can act to reduce drastically the practical efficiency as contrasted with the E_{\max} efficiency. These include the following:

1 The quantum yields taken from the literature are often measured under optimum laboratory conditions and usually represent maximum yields. (Often photolyses are carried to less than one per cent decomposition.)

2 Interference by the photolysis products (dissociation, light absorption, etc.) sometimes wastes solar energy.

3 Reverse reactions, which are exothermic, may be rapid. Thus "photostationary states" are reached where overall concentrations of products reactants remain unchanged. Comparatively little is known about these photochemical equilibria.

4 An appreciable amount of energy may have to be used

to operate the photolysis–fuel cell system, for gas pumping, product separation, etc.

It is well to recognize the limitations in photochemical efficiency due to these and other causes, as well as the limitations inherent in the sun as a source. However, the possibilities of large-scale photogalvanic devices and of energy storable photoregenerative fuel cells are attractive, nevertheless. Many of the practical difficulties will undoubtedly be solved as research in this field progresses.

Photoregenerative Fuel Cells

In a photoregenerative fuel cell system the vital endothermic photolysis step produces components at different oxidation states than originally present in the absorbing material. A critical factor is the separation and storage of these energy-rich components. Since their back reaction is exothermic, the rate of their separation has to be faster than their rate of spontaneous recombination, which gives only thermal energy. After separation, the exothermic recombination is brought about indirectly via oxidation and reduction reactions of the photolysis products at the electrodes. The components are returned to the oxidation state of the starting material and electrical power is produced. The system thus requires suitable electrodes and electrolyte so that the electrode reactions, the reformation of the starting material from the products of the electrode reactions, and the conductance of current are all efficient processes. There also can be no competing irreversible reactions from either the photolysis or the fuel cell steps, or else the cyclic nature of the process is impaired.

The difficulty in finding a fuel cell process which is suitable for photoregeneration can be appreciated by examining the list of interesting fuel cell components in Table 2. Although reaction 26 has a V^0 of only 0.21 V, we consider it as the most promising for this purpose. There are a number of different reasons, many of which have been mentioned in our general discussion, why the other twenty-five reactions are considered less promising. Included are:

1 The lack of absorption of solar energy by the fuel cell product, or greater absorption by the photolysis products than by the fuel cell product.

2 Small quantum yields.

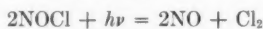
3 Undesirable side reactions, or irreversible reactions.

4 Lack of a convenient electrolyte.

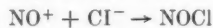
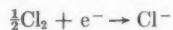
5 Disadvantage of a solid phase system.

The best known fuel cell is the much studied hydrox-type in which hydrogen and oxygen are reacted at the electrodes, using a hydroxyl ion electrolyte, with the resultant formation of water (reaction 11). Good efficiencies and high current densities have been reported on experimental models. The stability of hydrogen-oxygen mixtures (under appropriate conditions) also add to the desirability of having a regenerative fuel cell based on the photodecomposition of water. Unfortunately, although the photosensitized decomposition of water with a ceric-cerous system and a ferric-ferrous system has been demonstrated to occur at 2537 Å, the amount of photolysis from the solar spectrum is negligible (11).

Recent research⁶ has shown that nitrosyl chloride is a suitable substance for study as a photoregenerative fuel cell. The regenerative photolysis step for NOCl vapor is



For the power production step, the electrode reactions can be written in a simplified version as;



or



⁶ From research at Sundstrand Corp., Pacoima Laboratories supported by WADD. Aeronautical Accessories Lab., Power Div.

Table 2 Maximum performance of various fuel cell reactants 25C

Reaction	V°, volt	-ΔF°, kcal	watt hr per lb total fuel
1 $\frac{1}{2}\text{H}_{2(g)} + \frac{1}{2}\text{F}_{2(g)} \rightarrow \text{HF}_{(aq)}$	2.87	66.1	1740
2 $\text{O}_{2(g)} + \text{H}_{2(g)} \rightarrow \text{O}_{2(g)} + \text{H}_2\text{O}_{(l)}$	2.07*	95.8	1010
3 $\text{ClO}_{2(g)} + \frac{1}{2}\text{H}_{2(g)} + \text{OH}^-(aq) \rightarrow \text{H}_2\text{O}_{(l)} + \text{ClO}_{2(aq)}$	1.99	45.9	282
4 $\text{H}_2\text{O}_{2(l)} + \text{H}_{2(g)} \rightarrow 2\text{H}_2\text{O}_{(l)}$	1.86	86.1	1240
5 $\text{ClO}_2^-(aq) + 3\text{H}_{2(g)} \rightarrow \text{Cl}^-(aq) + 3\text{H}_2\text{O}_{(l)}$	1.72	79.1	780
6 $\text{HClO}_{2(aq)} + \text{H}_{2(g)} \rightarrow \text{HClO}_{(aq)} + \text{H}_2\text{O}_{(l)}$	1.64	75.9	473
7 $\text{HBrO}_{(aq)} + \frac{1}{2}\text{H}_{2(g)} \rightarrow \frac{1}{2}\text{Br}_{2(aq)} + \text{H}_2\text{O}_{(l)}$	1.55	35.8	193
8 $\text{BrO}_2^-(aq) + 3\text{H}_{2(g)} \rightarrow \text{Br}^-(aq) + 3\text{H}_2\text{O}_{(l)}$	1.44	199.7	784
9 $\frac{1}{2}\text{H}_{2(g)} + \frac{1}{2}\text{Cl}_{2(aq)} \rightarrow \text{HCl}_{(aq)}$	1.43	33.0	476
10 $\text{CO}_{(g)} + \frac{1}{2}\text{O}_{2(g)} \rightarrow \text{CO}_{2(aq)}$	1.29	59.5	712
11 $\text{H}_{2(g)} + \frac{1}{2}\text{O}_{2(g)} \rightarrow \text{H}_2\text{O}_{(l)}$	1.23	56.7	1660
12 $\text{HClO}_{2(aq)} + \text{H}_{2(g)} \rightarrow \text{HClO}_{2(aq)} + \text{H}_2\text{O}_{(l)}$	1.21	56.0	341
13 $\text{HClO}_{4(aq)} + \text{H}_{2(g)} \rightarrow \text{HClO}_{4(aq)} + \text{H}_2\text{O}_{(l)}$	1.19	54.8	282
14 $\text{Ag}_{(s)} + \frac{1}{2}\text{Cl}_{2(g)} \rightarrow \text{AgCl}_{(s)}$	1.13	26.2	96.1
15 $\frac{1}{2}\text{H}_{2(g)} + \frac{1}{2}\text{Br}_{2(aq)} \rightarrow \text{HBr}_{(aq)}$	1.11	25.6	168
16 $\text{CH}_4(g) + 2\text{O}_{2(g)} \rightarrow \text{CO}_{2(g)} + 2\text{H}_2\text{O}_{(l)}$	1.06	195.5	1280
17 $\text{N}_2\text{O}_{4(g)} + \text{H}_{2(g)} \rightarrow 2\text{HNO}_{2(aq)}$	1.06	49.1	275
18 $\text{H}_2\text{SO}_{3(aq)} + \frac{1}{2}\text{O}_{2(g)} \rightarrow \text{H}_2\text{SO}_{4(aq)}$	1.05	48.8	262
19 $\text{Ag}_{(s)} + \frac{1}{2}\text{Br}_{2(l)} \rightarrow \text{AgBr}_{(s)}$	1.01	23.4	65.6
20 $\text{Ag}_{(s)} + \text{HCl}_{(aq)} + \frac{1}{2}\text{O}_{2(g)} \rightarrow \text{AgCl}_{(s)} + \frac{1}{2}\text{H}_2\text{O}_{(l)}$	1.00	23.2	89.1
21 $\text{Fe}^{2+}(aq) + \frac{1}{2}\text{H}_{2(g)} \rightarrow \text{Fe}^{3+}(aq) + \text{H}^+(aq)$	0.77	17.77	164
22 $2\text{Ag}_{(s)} + \text{I}_{3(aq)} \rightarrow 2\text{AgI}_{(s)} + \text{I}^-(aq)$	0.68	31.7	28.0
23 $\text{H}_{2(g)} + \text{I}_2^-(aq) \rightarrow 2\text{HI}_{(aq)} + \text{I}^-(aq)$	0.53	24.7	51.0
24 $\text{CO}_{(g)} + \text{Cl}_{2(g)} \rightarrow \text{COCl}_{2(g)}$	0.38	17.5	96.1
25 $2\text{Fe}^{2+}(aq) + 3\text{I}_{(aq)} \rightarrow 2\text{Fe}^{3+}(aq) + \text{I}_3^-(aq)$	0.23	10.8	11.5
26 $\text{NO}_{(g)} + \frac{1}{2}\text{Cl}_{2(g)} \rightarrow \text{NOCl}_{(g)}$	0.21	4.86	39.1

Subscripts: g = gas; l = a liquid; s = a solid; aq = an aqueous solution

Metal chlorides such as aluminum chloride dissolve in NOCl to form the electrolyte (12), which in the unsolvated form may be represented as



Thus, the substance being photolyzed also forms the ionizing solvent for the electrolyte, considerably simplifying this system. A fuel cell of this type has been developed, using AlCl_3 in NOCl as an electrolyte and using NO and Cl_2 as the gaseous fuels.

The relatively fast spontaneous back reaction of NO with Cl_2 is indicative of one of the major problems in the development of a photoregenerative fuel cell. This is illustrated in Fig. 1 which shows that while 10 min of sunlight caused about 80% decomposition of a NOCl_(g) sample, the rate of the back reaction at this pressure is such that half of the photolysis products recombined after about 15 min in the dark. The data in Fig. 1 are replotted in Fig. 2 to show that the rate of the back reaction is given by the third-order rate equation

$$\frac{d(\text{NOCl})}{dt} = k_2 (\text{NO})^2 (\text{Cl}_2)$$

Photogalvanic Cells

Photogalvanic cells generally arise from an electron transfer spectrum of the light absorbing species, causing photooxidation-reduction as an electrolyte near an electrode (or of an electrode in an electrolyte) is illuminated. The electrical potential includes such effects as the concentration ratio of oxidized and reduced forms, and the relative rates of reaction of these species at the electrodes. Although it would be desirable to classify photogalvanic cells by the mechanisms of their photochemical and electrode processes, most of these processes are not yet well enough understood to make this possible. In reviews of the subject, classification has been

made partly on the basis of the physical relationships of the electrode and the absorbing species. Four types have been described (3).

- Type 1 Metal electrodes immersed in solutions of electrolytes.
- Type 2 Metal electrodes coated with inorganic compounds and immersed in solutions of electrolytes.
- Type 3 Metal electrolytes coated with a dye and immersed in solutions of electrolytes.
- Type 4 Metal electrodes immersed in organic liquids.

While the great advantage of photogalvanic cells lies in the possibility of economical, direct conversion of light into electrical power, there are many problems that must be studied regarding the fundamentals of these processes before photogalvanic systems can best be utilized. The maximum benefit from research studies can be expected to come from intensive studies on a particular system in which as much information as possible can be correlated to reach an understanding of the variables in the system. The effect of variables such as concentration, light intensity, temperature, pH, etc., on the following aspects of photogalvanic cells (and photoregenerative fuel cells) are those that we consider to be the most important.

- 1 The mechanism of the photochemical primary process and its quantum efficiency.
- 2 The rate of the spontaneous back reactions that compete with the electrode reactions, and means of inhibiting them.
- 3 The mechanisms of the electrode reactions. The extent and cause of electrode polarization.
- 4 The design of illumination geometry and electrode positions with regard to these mechanisms and to the diffusion of species to and from the electrodes.
- 5 The use of sensitizers to extend the range of efficient photochemical reaction.

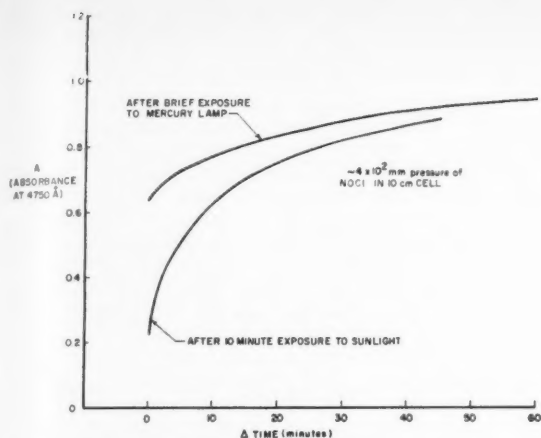


Fig. 1 Back reaction from NOCl photolysis, $2\text{NO}_{(g)} + \text{Cl}_{2(g)} = 2\text{NOCl}_{(g)}$. (The absorbance of the sample before exposure and after a long period in the dark = $A_{\infty} = 1.186$)

Many systems have been reported (or might be expected) to have photogalvanic properties, but detailed studies have been reported on only a few of these. The best known of these is the thionine-iron system, studied originally by Rabinowitch (13). The blue dye thionine absorbs light, is bleached as ferrous ions are oxidized to ferric ions, and produces a potential of about 0.25 v. The overall reaction can be represented as

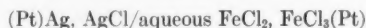


Rabinowitch has indicated (14) that the thionine-iron system has produced voltages up to 0.5 v, and has converted about 1% of the absorbed light into electrical energy without optimization of conditions. The much poorer results reported in recent work by Potter and Thaller (15) may be in conflict with those of Rabinowitch. However, some of the conditions used in their studies such as the use of chloride as the anion instead of sulfate might be expected to give lower efficiencies.

Some interesting work has been done on electrodes coated with inorganic compounds, such as the copper oxide system (3)



and the silver halide system (16)



With the use of an anion-impermeable membrane, and phosphate complexing of ferric ions, the latter has been considered as a photogalvanic battery for the direct storage of solar energy. However, it should be noted that if a voltage of 50 mv is assumed, then with the largest useable thickness reported for AgCl (20 μ) can store a maximum theoretical power of only about 1×10^{-4} w-hr/cm² after even 100% conversion to Ag. This is about 1000 times less than that actually delivered by lead storage batteries (17).

The direct storage of electrochemical energy in the photogalvanic cell itself means that it must act essentially as a battery to be able to deliver current in the dark. There seems no a priori reason to believe that a photogalvanic battery would have characteristics superior to electrically regenerated storage batteries. On the contrary, the photogalvanic batteries which have been proposed have the same type of phase, temperature and stability disadvantages as

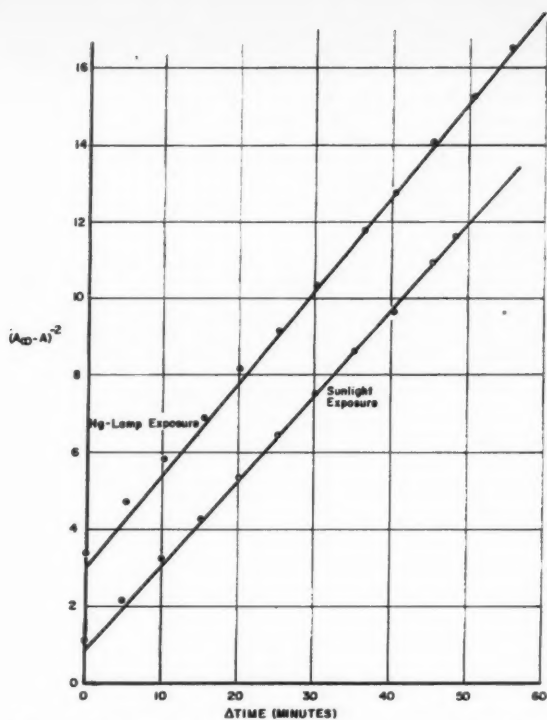


Fig. 2 Third order plot of photolysis back reaction at 17°C, $2\text{NO}_{(g)} + \text{Cl}_{2(g)} = 2\text{NOCl}_{(g)}$

other storage batteries and suffer from other limitations as well. Firstly, the thermal back reaction of the photolysis products usually competes with the desired electrode reaction. Storage requires inhibition of this back reaction. (In the ordinary storage battery the reactive components are at least separated at the electrodes in the charging process.) Secondly, optimum operation of photogalvanic cells generally involves the use of a small amount of photoactive substance with a high coefficient of light absorption. The amount of this photoactive substance per volume or per electrode area is quite small, and thus yields very little direct storage of electrochemical energy. Thirdly, the photopotentials of nearly all photogalvanic cells are in the 0.05 to 0.5 v range, which alone means that direct storage of the electrochemical energy will require photogalvanic batteries at least 4 to 40 times as large as the 2 v per cell storage batteries.

In general, the direct storage of electrochemical energy is more promising in photoregenerative fuel cells rather than in photogalvanic batteries. The best utilization of photogalvanic cells appears to be in either the immediate use of the electrical energy or else its storage in conventional storage batteries. (Another means of storage is the regeneration of fuel cell reactants by electrolysis.)

Conclusions

In theory, photochemical processes represent possible means of converting solar energy to useful power, both on the Earth's surface and in space vehicles. However, devices such as solar regenerative fuel cells and photogalvanic cells are still in the early developmental stages. Significant advances have been made recently, but considerable research and development on the fundamental and applied photo-

chemistry of chemical systems must be expended if the true potential of these important devices is to be realized. Both types of these photochemical systems should be adaptable to large scale power conversion units using simple solar concentrators. The photoregenerative fuel cell type of device is of most interest for the direct storage of electrochemical energy.

References

- 1 (a) Calvert, J. G., "Photochemical Processes for Utilization of Solar Energy," *Ohio J. Sci.*, vol. 53, p. 293, 1953. (b) Calvert, J. G., *Introduction to Solar Energy*, A. M. Zarem, editor, McGraw-Hill Co., N. Y., in press.
- 2 Rosenberg, N., "Photochemical Processes and their Utilization for Solar Energy," "Proceedings of a Seminar on Advanced Energy Sources and Conversion Techniques," ASTIA no. AD 209301 vol. 1, p. 139, 1958.
- 3 Sancier, K. M., "Photogalvanic Cells," *Transactions of the Conference on the Use of Solar Energy—The Scientific Basis*, vol. 5, pp. 43-51, 1955.
- 4 Zwick, E. B. and Zimmerman, R. L., "Space Vehicle Power Systems," *ARS JOURNAL*, vol. 29, pp. 553-564, 1959.
- 5 Friedman, D., *NRL Memorandum Rep.* 1005, Dec. 1959.
- 6 Fritz, S., "Compendium of Meteorology," T. F. Malone, editor, American Meteorological Soc., Boston, Mass., p. 13, 1951.
- 7 Kistiakowsky, G. B., "Photochemical Decomposition of Nitrosyl Chloride," *J. Amer. Chem. Soc.*, vol. 52, p. 102, 1930.

- 8 Bailey, C. R. and Cassie, A. B. D., "Investigations in the Infra-red Region of the Spectrum X. The Asymmetric Molecule Nitrosyl Chloride, NOCl," *Proc. Roy. Soc., London, Eng.*, A 145, pp. 336-344, 1934.
- 9 Goodeve, C. F. and Katz, S., "The Absorption Spectrum of Nitrosyl Chloride," 1939, *Proc. Roy. Soc., London, England*, A 172, p. 432.
- 10 Natanson, G. L., "On the Mechanism of Photodecomposition of NOCl," *Acta Physicochimica URSS*, vol. 11, pp. 521-536, 1939.
- 11 Heidt, L. J. and Smith, M. E., "Influence of Perchloric Acid and Cerous Perchlorate upon the Photochemical Oxidation of Cerous to Ceric Perchlorate in Dilute Aqueous Perchloric Acid," *J. Amer. Chem. Soc.*, vol. 70, p. 2476, 1948, see also L. J. Heidt, and A. F. McMillan, *ibid.*, vol. 76, p. 2135, 1954 J. S. Forcyce, Ph.D. Thesis in Physical Chemistry, MIT, Oct., 1958.
- 12 (a) Burg, A. B. and Campbell, G. W. Jr., "Liquid Nitrosyl Chloride as an Ionizing Solvent," *J. Amer. Chem. Soc.*, vol. 70, pp. 1964-1965, 1948. (b) Burg, A. B. and McKenzie, D. E., "Further Studies of Nitrosyl Chloride as an Ionizing Solvent," *J. Amer. Chem. Soc.*, vol. 74, pp. 3143-3147, 1952.
- 13 Rabinowitch, E., "The Photogalvanic Effect I. The Photochemical Properties of the Thionine-Iron System II. The Photogalvanic Properties of the Thionine-Iron System," *J. Chem. Phys.*, vol. 8, pp. 551-566, 1940.
- 14 Rabinowitch, E., "Photochemical Utilization of Light Energy," "Solar Energy Research," F. Daniels and J. A. Duffie, editors, U. of Wisconsin Press, Madison, Wis., pp. 200-202, 1955.
- 15 Potter, A. E., Jr. and Thaller, L. H., "Efficiency of Some Iron-Thionine Photogalvanic Cells," *J. Solar Energy Science & Engng.*, vol. 3, pp. 1-7, 1959.
- 16 Lasser, M. E., Zaromb, S. and Kalhammer, F., "Survey of Photogalvanic Systems," *Proc. 13th Annual Power Sources Conf.* (Power Sources Div., U. S. Army Signal R & D Lab, Fort Monmouth, N. J.) April 28-30, 1959.
- 17 Vinal, G. W., "Storage Batteries," 4th edition, John Wiley & Sons, N. Y., pp. 213-215, 1955.

Internal Design Consideration for Cavity-Type Solar Absorbers

CHARLES W. STEPHENS¹
and ALAN M. HAIRE²

Electro-Optical Systems, Inc.
Pasadena, Calif.

The effects of geometry and internal surface reflectivity on the performance of a cavity-type absorber used in solar power systems are discussed. By defining an average "view factor" of the cavity opening as seen by the cavity interior surface, it is possible to evaluate quantitatively the performance of a cavity operating at elevated temperatures. Only minor performance variations between the several cavity shapes considered existed, and it was found possible to approximate closely the ideal blackbody cavity absorber efficiency with real cavities of reasonable size. The analysis demonstrates that high performance cavities are attainable with internal surface reflectances on the order of 0.6 to 0.7. For these values, differences in heat absorption rate throughout the cavity may be minimized with the consequent reduction of the problem of internal "hot spots."

AN IMPORTANT class of space power systems which will be utilized for the next several years in the lower power ranges (less than 15-25 kw) will be the solar-thermal converter. The list of problem areas associated with the various solar-thermal converters that might be envisioned is long, but there are certain areas which are of common interest to all converters due to the nature of the energy source. These

common areas are the concentrator selection, accuracy, reflective surface characteristics, and orientation; also the absorber selection and design. This paper will discuss the absorber selection and design, and the effect upon absorber performance of the various shape factors, relative absorber size, temperature and spectral emissivity characteristics.

Selection of Absorber Type

The three most promising absorber configurations are the cavity, the hemisphere and the flat plate. As will be shown later, a cavity-type absorber (shown schematically in Fig. 1) can be made to closely approach the radiation characteristics

¹ Presented at the ARS Semi-Annual Meeting, Los Angeles, Calif., May 9-12, 1960.

² Associate Manager, Advanced Power Systems Div. Member ARS.

³ Engineer, Advanced Power Systems Div.; presently, Project Engineer, Aerolab Development Project Co., Pasadena, Calif.

of a black body, with an effective emissivity over all wavelengths approaching 1 at the cavity entrance. The hemisphere and flat plate appear promising as solar absorber configurations when the use of a spectrally selective surface is considered. This spectrally selective surface would be one that would have a high absorptance to solar wavelengths and a low absorptance to longer wavelengths, corresponding to reradiation at the absorber temperature. Fig. 2 shows a comparison between the cavity, hemisphere and flat plate. In this figure, a perfect cavity is compared with a hemisphere and flat plate having extremely optimistic spectrally selective characteristics. Surfaces with these spectral characteristics do not, to the authors' knowledge, exist at present, although, with the expanded activities in the field of spectral selectivity, it is not unreasonable to assume that such surfaces might be developed within the next three years. It should be noted in Fig. 2 that there is an intrinsic loss with the hemisphere and the flat plate which is equal to the reflectance of the absorber surface to wavelengths characteristic of solar radiation. The compensating factor, however, is the lower value of energy reradiated, as is obvious from the lesser slope of the hemisphere and flat plate absorber loss curves as compared with the cavity. At higher temperatures, the crossover point will occur at smaller ratios of the cavity diameter to the sun's image diameter. (The magnitude of this ratio required to encompass the 80% of the total flux incident on the mirror is a function of the concentrator surface and orientation accuracies.) See (5). For a power system requiring temperatures usually associated with thermionic power supplies (on the order of 2200 F) with mirrors of reasonable surface and orientation accuracy, it appears that a spectrally selective flat plate or hemispherical absorber with the spectral characteristics delineated in Fig. 2 would display lower energy losses.

An important advantage which occurs from the use of the cavity absorber is the considerable design freedom allowed from the standpoint of heat transfer area for removing energy from the absorber. There are no intrinsic limitations on cavity size so long as it satisfies the requirement of sufficient reflections to absorb all but a small fraction of the incident solar radiation, as will be discussed later. By lengthening, widening or otherwise altering the internal geometry of the cavity, it is possible, therefore, to utilize the configuration most optimum from the standpoint of the design requirements. With the hemisphere and flat plate, this freedom does not exist and heat transfer coefficients on the order of 2000 w/in.² will be necessary to remove the heat from the small areas available.

The Cavity Absorber and Its Practical Approximations

In nature, an exactly black surface does not exist. For a number of years, scientific experiments requiring essentially "black" radiation have approached the problem in the following manner. A hollow space is used in which there is a small opening (a hohlraum) with the walls of the space kept at uniform temperature. Radiation which streams through the opening into the hollow space (see Fig. 1) is reflected repeatedly before it leaves again through the opening. Depending upon the surface absorptivity, after a number of reflections the percentage of radiation absorbed will closely approach 100%. This means that the opening of the hollow space acts like a black surface, since all wavelengths are totally absorbed. Accordingly, all wavelengths must also be totally emitted according to Kirchhoff's Law.

For an ideal cavity, the view factor θ from the cavity interior to the opening, is equal to zero, which means that no radiation entering the cavity is reflected back out of the cavity. It would be impossible to determine the location of any radiation emitted by the cavity regarding the internal portion of the cavity from which it emanated. Due to the existence of the hohlraum, the performance of the cavity is

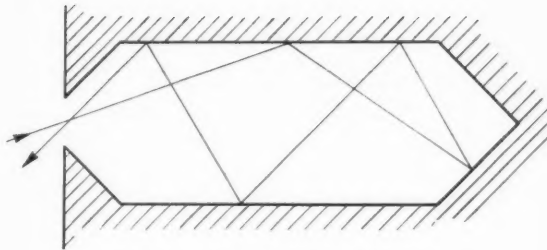


Fig. 1 Cavity absorber schematic

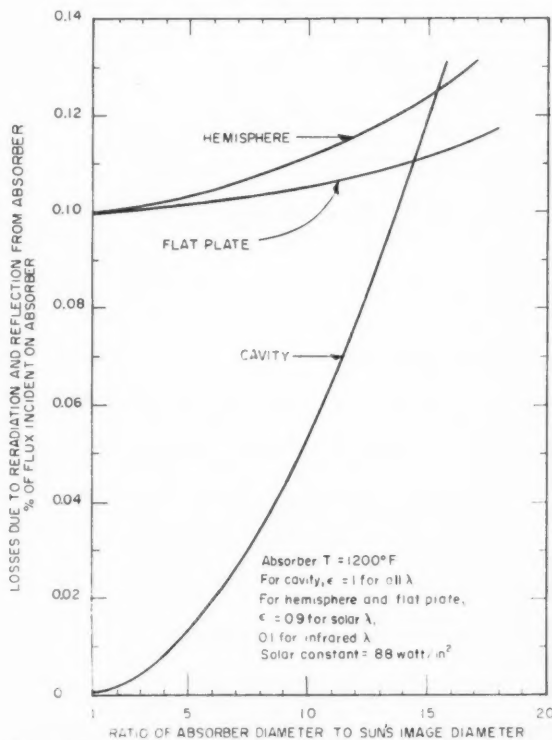


Fig. 2 Per cent loss vs. absorber size for various absorbers. Assumes that 80% of flux incident on mirror will fall on the absorber, independent of absorber diameter

totally unaffected by the internal emissivity, as long as it does not equal zero.

In designing a real cavity, the view factor θ , the internal surface emissivity, and the cavity geometrical configuration must be considered in providing a cavity absorber design which absorbs the maximum net amount of the energy incident upon it with a reasonable total weight. The cavity should also provide a reasonably even distribution of energy over its internal surface to minimize local hot spots and to improve the general heat exchange characteristics between the cavity and the thermal converter working media.

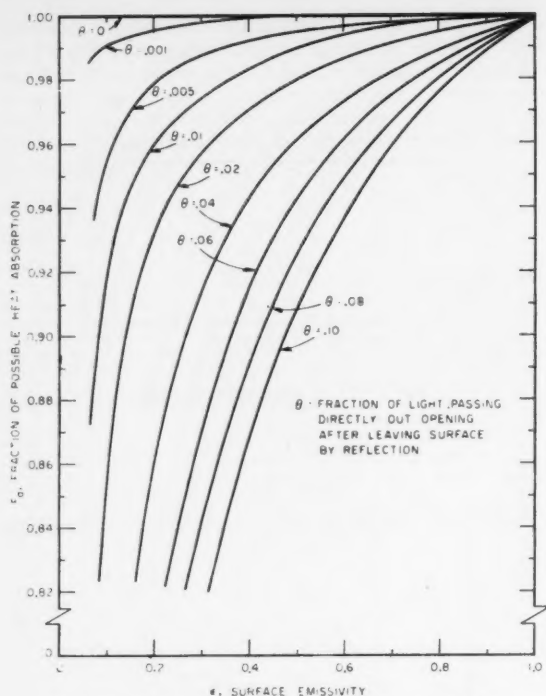


Fig. 3 Cavity heat absorption vs. surface emissivity

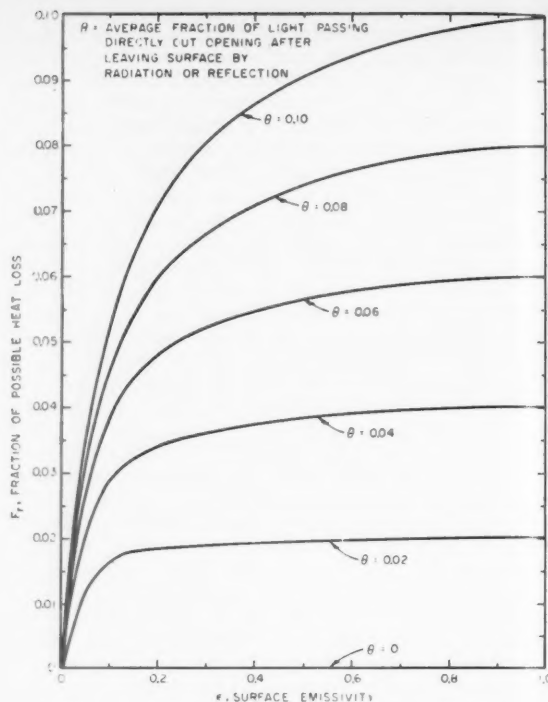


Fig. 4 Cavity heat loss vs. surface emissivity

Cavity Design Criteria

For the purposes of this study, the following assumptions will be made:

1 The internal cavity surface temperature is maintained at a constant value. Since most of the solar-thermal power systems envisioned to date utilize some type of isothermal phase-change thermal storage, the assumption of essentially constant cavity temperature appears quite reasonable.

2 The amount of flux entering the cavity, per unit area of cavity opening A_0 is constant. This factor is fixed by the concentrator selection, accuracy, reflection surface characteristics, and orientation; these should not be affected by the absorber in any manner.

3 An average view factor θ may be defined such that it represents the average fraction of light which is either reflected or emitted from the surface of the cavity interior and which passes directly out of the cavity through the cavity opening. All radiation reflected by a surface is assumed to be hemispherically diffuse with an intensity varying as the cosine of the angle with the surface normal.

When solar energy enters the cavity entrance, as shown in Fig. 1, a certain fraction of the light reflected on each successive reflection will pass directly through the cavity opening. The total amount of energy absorbed on the cavity interior surface may be analyzed as follows:

At the first reflection $Q_i\epsilon$ is absorbed and $Q_i(1 - \epsilon)$ is reflected. Of the reflected light

$$Q_i(1 - \epsilon)(\theta)$$

passes through the opening and

$$Q_i(1 - \epsilon)(1 - \theta)$$

is reflected again. Of this light

$$Q_i(1 - \epsilon)(1 - \theta)(\epsilon)$$

is absorbed and

$$Q_i(1 - \epsilon)(1 - \theta)(1 - \epsilon)$$

is reflected. Following the same reasoning the light absorbed on the n th reflection after the first is

$$Q_i(1 - \epsilon)^n(1 - \theta)^n\epsilon$$

The total heat absorbed is

$$Q_i\epsilon + \sum_{n=1}^{\infty} Q_i(1 - \epsilon)^n(1 - \theta)^n\epsilon$$

or

$$Q_i\epsilon \left[1 + \sum_{n=1}^{\infty} (1 - \epsilon)^n(1 - \theta)^n \right]$$

or

$$Q_i\epsilon \left[\frac{1}{1 - (1 - \epsilon)(1 - \theta)} \right]$$

The fraction of possible heat absorbed F_a is

$$F_a = \frac{\epsilon}{1 - (1 - \epsilon)(1 - \theta)}$$

This fraction is plotted in Fig. 3.

It should be noted in Fig. 3 that for values of θ less than 0.01, the change of F_a is extremely slight for values of surface emissivity between 0.3 and 1.0. In this range, therefore, the amount of energy lost by reflection out of the cavity entrance is relatively insensitive to changes in emissivity above values of 0.3.

The total amount of energy lost through the cavity opening by reradiation from the cavity interior surface may be analyzed similarly:

$Q_b \epsilon \theta$ passes directly through the opening and $Q_b \epsilon (1 - \theta)$ strikes the cavity surface. Of this

$$Q_b \epsilon (1 - \theta) \epsilon$$

is absorbed and

$$Q_b \epsilon (1 - \theta) (1 - \epsilon)$$

is reflected. Of the reflected light

$$Q_b \epsilon (1 - \theta) (1 - \epsilon) (1 - \theta)$$

again strikes the cavity surface. Following this same reasoning the heat leaving the cavity after the n th reflection is

$$Q_b \epsilon (1 - \theta)^n (1 - \epsilon)^n \theta$$

The total heat leaving the cavity is

$$Q_b \epsilon \theta + \left[\sum_{n=1}^{\infty} Q_b \epsilon (1 - \theta)^n (1 - \epsilon)^n \theta \right]$$

or

$$Q_b \epsilon \theta \left[1 + \sum_{n=1}^{\infty} (1 - \theta)^n (1 - \epsilon)^n \right]$$

$$Q_b \epsilon \theta \frac{1}{1 - (1 - \theta)(1 - \epsilon)}$$

The fraction of possible heat lost F_r is

$$F_r = \epsilon \theta / [1 - (1 - \theta)(1 - \epsilon)]$$

This fraction is plotted in Fig. 4. It is interesting to note in Fig. 4 that for values of θ less than 0.06 that the fraction of possible heat loss F_r is quite insensitive to variations in emissivity between 0.3 and 1.0. The emissivities delineated in Fig. 3 are averaged through the solar wavelengths, approximately 0.3 to 2 μ . The emissivities of Fig. 4 are those corresponding to the reradiation spectrum of a black body at temperature of the cavity (between 2 and 10 μ).

The general expression for θ can be derived for a variety of shapes in terms of their geometrical characteristics as well as the ratio of cavity entrance area A_o and the cavity interior surface area A_c . Assuming that the density of radiation leaving any portion of the surface of the cavity interior is proportional to the cosine of the angle with the surface normal (totally diffuse or hemispherical emission or reflection), θ is given by the following expression

$$\theta = \frac{1}{\pi} \frac{\int \gamma \cos \psi dA_c}{A_c} \quad [1]$$

where the integral is taken over the cavity interior. Utilizing this expression, the cavity shapes shown in Fig. 5 were evaluated in terms of their geometrical parameters and A_o/A_c . These shapes are symmetrical about the axis of the cavity entrance.

It was not possible to obtain the integral of Eq. 1 analytically for the various shapes considered. Numerical integration, however, yielded the curves of θ vs. A_o/A_c shown in Fig. 6. Since, in order to minimize both losses by direct reflection as well as reradiation, it is desirable to minimize θ , the sphere appears to be a most desirable cavity shape for values of A_o/A_c greater than approximately 0.025. It is interesting to note, however, that, except for the hemisphere, there appears to be no considerable variation in the value of θ between the shapes considered.

By combining the results shown in Figs. 3, 4 and 6, it is now possible to ascertain the performance characteristics of the various cavity shapes as follows:

The amount of heat absorbed is given by

$$Q_a = F_a C Q A_o \quad [2]$$

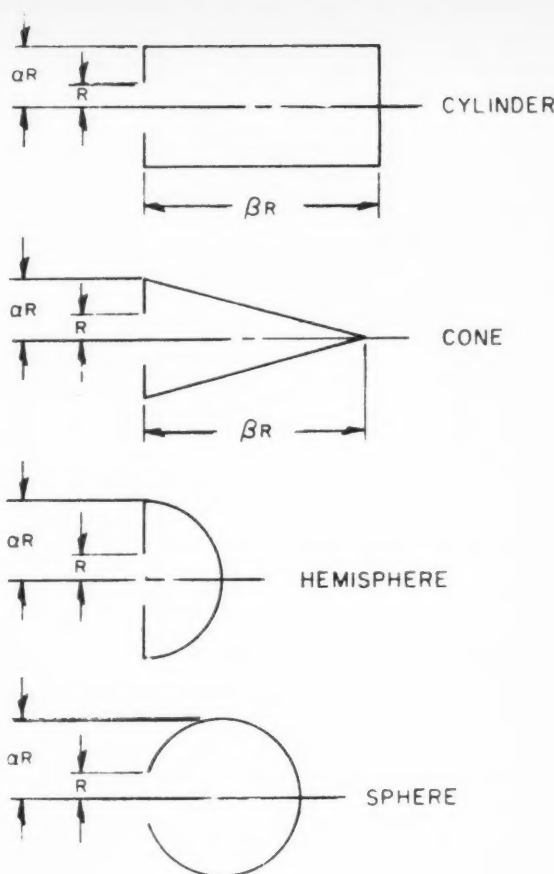


Fig. 5 Typical cavity shapes

The amount of heat lost by reradiation is

$$Q_r = F_r \sigma T^4 A_c \quad [3]$$

The efficiency or fraction of possible heat absorbed is

$$\eta = \frac{F_a C Q A_o - F_r \sigma T^4 A_c}{C Q A_o} \quad [4]$$

$$\eta = F_a - F_r \frac{\sigma T^4}{C Q} (A_c/A_o) \quad [5]$$

By selecting a value of internal emissivity, F_a and F_r may be obtained from Figs. 3 and 4, respectively. C may be determined from the characteristics of the solar concentration system, and curves of η vs. A_o/A_c may be calculated for various values of temperature.

Fig. 7 shows the results of these calculations for the spherical absorber with a concentration ratio of 1250, which is a typical value for a paraboloid mirror. The value of emissivity utilized for all wavelengths was 0.3. It should be noted that the values shown in these curves would be little changed for values of θ less than approximately 0.02 and for emissivities greater than 0.3 because of the very slight variation of F_a and F_r with emissivity in this range of θ , as mentioned previously. The efficiency values shown for $A_o/A_c = 0$ correspond to those obtainable with a true blackbody cavity. For this case, the cavity efficiency expression reduces to the following simplified form

$$\eta = 1 - \sigma T^4 / C Q \quad [6]$$

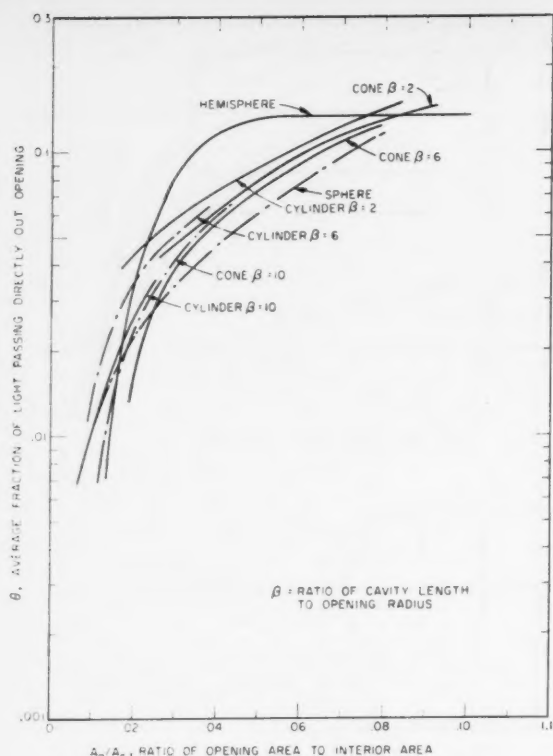


Fig. 6 Light passing out of opening of various cavities

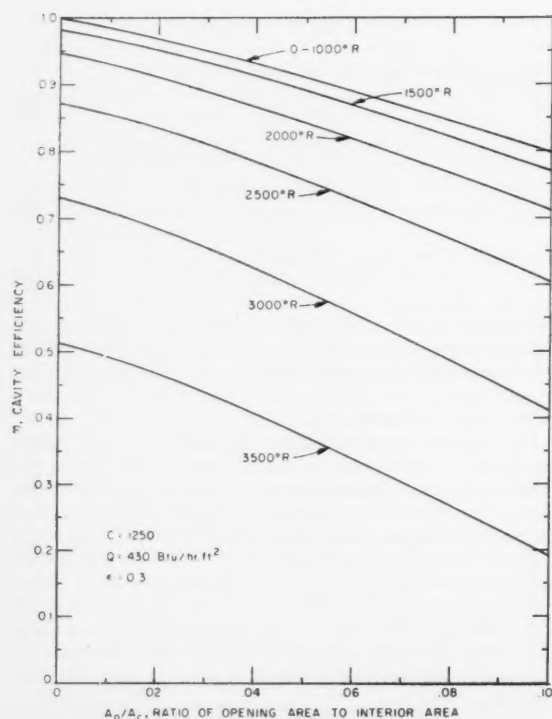


Fig. 7 Efficiency of spherical cavity at various temperatures

It would be generally difficult, if not impossible, to achieve exactly uniform illumination of the cavity by the concentrated solar radiation passing through the cavity entrance. It is generally desirable, however, that the energy absorption of any given elemental area of the cavity interior surface be as nearly constant as possible to prevent the formation of hot spots and to allow the utilization of an approximately constant thickness layer of surrounding thermal storage material. The interior emissivity is the governing factor in the distribution of energy absorption. Its effects may be studied by the use of a simple model. A cavity consisting of two identical sections will be considered. The two sections receive light of different intensity, but they reflect light identically. The net distribution can be analyzed as follows:

Considering section 1, at the first reflection Q is absorbed and $Q_1(1 - \epsilon)$ is reflected. Of the reflected light

$$Q_1(1 - \epsilon)(\theta)$$

passes through the opening and

$$Q_1(1 - \epsilon)(1 - \theta)$$

strikes section 2. Of this light

$$Q_1(1 - \epsilon)(1 - \theta)\epsilon$$

is absorbed by section 2 and

$$Q_1(1 - \epsilon)(1 - \theta)(1 - \epsilon)$$

is reflected. Following this reasoning the light absorbed by surface 1 is

$$Q_1\epsilon + Q_2(1 - \epsilon)(1 - \theta)\epsilon + Q_1(1 - \epsilon)^2(1 - \theta)^2\epsilon + \dots$$

and by surface 2 is

$$Q_2\epsilon + Q_1(1 - \epsilon)(1 - \theta)\epsilon + Q_2(1 - \epsilon)^2(1 - \theta)^2\epsilon + \dots$$

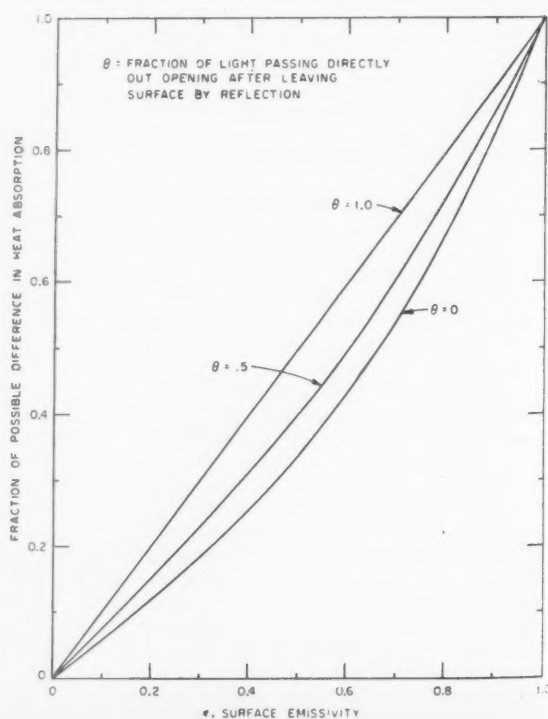


Fig. 8 Cavity heat distribution vs. surface emissivity

These can be reduced to

$$Q_1 \epsilon \frac{1}{1 - (1 - \epsilon)^2(1 - \theta)^2} + Q_2 \epsilon \frac{(1 - \epsilon)(1 - \theta)}{1 - (1 - \epsilon)^2(1 - \theta)^2}$$

for surface 1, and

$$Q_2 \epsilon \frac{1}{1 - (1 - \epsilon)^2(1 - \theta)^2} + Q_1 \epsilon \frac{(1 - \epsilon)(1 - \theta)}{1 - (1 - \epsilon)^2(1 - \theta)^2}$$

for surface 2. The difference in heat absorbed by the surfaces is

$$Q_1 \epsilon \left[\frac{1}{1 - (1 - \epsilon)^2(1 - \theta)^2} - \frac{(1 - \epsilon)(1 - \theta)}{1 - (1 - \epsilon)^2(1 - \theta)^2} \right] - Q_2 \epsilon \left[\frac{1}{1 - (1 - \epsilon)^2(1 - \theta)^2} - \frac{(1 - \epsilon)(1 - \theta)}{1 - (1 - \epsilon)^2(1 - \theta)^2} \right] =$$

$$(Q_1 - Q_2) \epsilon \left[\frac{1 - (1 - \epsilon)(1 - \theta)}{1 - (1 - \epsilon)^2(1 - \theta)^2} \right] = (Q_1 - Q_2) \epsilon \frac{1}{1 + (1 - \epsilon)(1 - \theta)}$$

Since $Q_1 - Q_2$ is the maximum possible difference in heat absorption, the fraction of possible difference in heat absorption is

$$\epsilon / [1 + (1 - \epsilon)(1 - \theta)]$$

This fraction is plotted in Fig. 8. It can be seen that it is desirable to have a low θ and a low surface emissivity (at solar wavelengths) to minimize the difference in heat absorption in the cavity. This is fortunately consistent with the requirement of a minimum θ to maximize the cavity efficiency.

Conclusions

For values of A_o/A_c on the order 0.03 or less, it is possible to closely approximate the blackbody cavity absorber efficiency. This efficiency is affected only slightly by variations in emissivity from 0.3 to 1.0 over all wavelengths.

The spherical absorber appears to be slightly more efficient than the other shapes considered for $A_o/A_c > 0.02$. For values of A_o/A_c on the order of 0.02 or less, little efficiency variations between shapes occur.

By utilizing a surface emissivity of 0.3-0.4 it is possible to minimize differences in heat absorption rates throughout the cavity. Since emissivities do not appear to seriously affect the cavity performance for A_o/A_c less than 0.03, the use of emissivities of this order are felt justified when it appears that thermal hot spots will be a problem.

For concentration ratios greater than that utilized in the study (1250) the effects on efficiency of lower emissivity, greater values of θ and larger values of A_o/A_c will be less than

those shown here. The effect on efficiency of these factors for concentration ratios less than 1250 will, however, be amplified.

Nomenclature

- A_c = area of cavity interior surface
- A_o = area of cavity opening
- C = concentration ratio of energy entering cavity
- F_a = fraction of energy entering cavity that is absorbed

- F_r = fraction of total energy emitted by cavity surface that passes out through the opening
- Q = solar constant in space, 430 Btu/ft² hr
- Q_a = total energy absorbed by cavity
- Q_b = energy emitted by a black body at temperature of cavity interior
- Q_t = total energy entering cavity, QCA_o
- R = cavity opening radius
- T = cavity temperature, R
- α = ratio of cavity height to opening radius (see Fig. 5)
- β = ratio of cavity length to opening radius (see Fig. 5)
- γ = solid angle subtended by cavity opening when viewed by element of cavity interior surface
- ϵ = emissivity of cavity interior
- σ = Stefan-Boltzmann Constant, 0.174×10^{-8} Btu/hr ft² R⁴
- θ = average view factor of cavity opening by cavity interior surface
- ψ = angle between line connecting elemental cavity surface with cavity opening and the surface normal
- η = cavity efficiency

References

- 1 Jose, P. D., "The Flux Through the Focal Spot of a Solar Furnace," *J. Solar Energy Sci. and Engng.*, Oct. 1957, p. 19.
- 2 Baum, W. A. and Strong, J. D., "Basic Optical Considerations in the Choice of a Design for a Solar Furnace," AFMDC TR58-7, *Solar Furnace Support Studies*, Vol. I, Holloman AFB, 1958, p. 57-88.
- 3 Silvern, D. H., "An Analysis of Mirror Accuracy Requirements for Solar Power Plants," ARS preprint 1179-60, May 1960.
- 4 Hiester, N. K., Tietz, T. E., Loh, E. and Duwez, P., "Theoretical Considerations on Performance Characteristics of Solar Furnaces," *JET PROPULSION*, vol. 27, no. 5, May 1957, p. 507.
- 5 McClelland, D. H., "Solar Concentrators for High Temperature Space Power Systems," ARS preprint 1314-60, Sept. 1960; see also "Design and Development of Solar Concentrators and Their Integration Into Space Power Systems," Electro-Optical Systems, Inc., Pasadena, Calif., Rep. no.410-1R-1, Nov. 20, 1959.

High Temperature Rarefied Hypersonic Flow Over a Flat Plate¹

H. T. NAGAMATSU²
R. E. SHEER JR.³
and J. R. SCHMID⁴

General Electric Co.
Schenectady, N. Y.

Aerodynamic investigations were conducted on the flow over a flat plate with a sharp leading edge in a hypersonic shock tunnel. The flow Mach number ranged from 7.95 to 25.1 with stagnation temperatures of approximately 2300 and 7000 R. For these flow conditions, the static pressure in the test section varied from 3.2μ to 0.95 mm of mercury and the Knudsen number, based on the ratio of the free stream mean free path and the leading edge thickness, varied from approximately 0.4 to 86.

Schlieren photographs taken at these conditions indicated the presence of slip flow at the sharp leading edge, in that the formation of the strong shock wave and the boundary layer was delayed. The induced pressures measured on the surface of the plate indicated the existence of slip flow in the vicinity of the leading edge. The slip flow region increased with the free stream flow Mach and Knudsen numbers.

In the strong shock wave-boundary layer interaction region near the leading edge, the viscous region extended from the plate surface to the shock wave, and there was no apparent inviscid region between the shock wave and the boundary layer. The induced pressure on the surface of the flat plate varied inversely as the square root of the distance from the leading edge for all cases except at the highest Mach number of 25.1. At this Mach number, the maximum induced pressure on the surface near the leading edge was 28 times the free stream static pressure. For a reflected stagnation temperature of 7000 R with real gas effects, the leading edge slip phenomenon and the shock wave-boundary layer interaction were present.

The strong interaction parameter $\chi = M_1^2/\sqrt{Re_x}$ varied from 2.5 to 756. The experimental induced pressures are compared with the insulated and noninsulated interaction theories of Li and Nagamatsu.

THE GROWING interest in glide-type hypersonic vehicles, terminally guided re-entry bodies, and space vehicles, has indicated that a need exists for fundamental knowledge regarding the aerodynamic characteristics of lifting surfaces in rarefied flow. This is especially true at the high flight Mach numbers and stagnation temperatures encountered during re-entry at high altitudes. For these hypersonic vehicles the flight Mach number ranges from about 15 to 35. With manned hypersonic vehicles, the deceleration and the control problem during re-entry becomes important; thus a glide-type vehicle appears most promising.

In hypersonic flow over a flat plate with a sharp leading edge, a shock wave and viscous boundary layer form at the leading edge. Due to the interaction of the leading edge shock wave and the viscous layer, the pressure over the plate surface is increased. At high flow Mach numbers, the induced pressure caused by this interaction can be many times greater than the free stream pressure. This increase in the surface pressure increases the skin friction as well as the heat transfer.

Recent experimental results (1-4) for a flat plate with

a sharp leading edge in a rarefied hypersonic flow have indicated that there exists a delay in the formation of the maximum shock wave angle and boundary layer at the leading edge. This delay in formation may be caused by the slip phenomena in the vicinity of the leading edge (1,3-5). After the slip region, the shock wave and the boundary layer were merged before separating at very high Mach numbers (4).

Both the formation and the delay of the maximum shock wave angle and the boundary layer are strongly influenced by the free stream Mach number, ratio of the mean free path to the leading edge thickness (Knudsen number), and the Reynolds number based upon the leading edge thickness. If the Knudsen number is very small, a detached shock wave forms in front of the plate and the flow after the normal shock wave is subsonic with relatively high pressure. For this condition, the viscous effects become small and the inviscid flow over the blunt leading edge predominates, as discussed in (6-9). Since the flow is subsonic ahead of the leading edge, there is no possibility of the existence of slip flow for a reasonably dense gas. For a flat plate with a sharp leading edge, Knudsen number equal to or greater than unity, a first order theory (5) on the slip phenomenon indicates that the slip distance is proportional to the flow Mach number, mean free path of the gas, and number of collisions. Hence, at very high flow Mach numbers the flow over a blunt body is drastically different than that for a body with a sharp leading edge.

For the region of strong interaction between the shock

Received July 10, 1960.

¹ This research was partially supported by the Ballistic Missile Division, U. S. Air Force, under contract AF 04-(647)-269.

² Research Associate. Also, Adjunct Professor, Rensselaer Polytechnic Institute, Troy, N. Y. Member ARS.

³ Gasdynamics, Research Lab.

⁴ Mechanical Engineer. Formerly with Missile and Space Vehicles Dept., General Electric Co., Philadelphia, Pa.

⁵ Numbers in parentheses indicate References at end of paper.

wave and boundary layer for a flat plate with infinitely thin leading edge, a number of theories have been developed for the case with no slip on the surface. They differ primarily in the assumption that the shock wave and boundary layer are either separated or merged. Recent schlieren photographs at high flow Mach numbers (4) and experiments using a free molecular probe in the vicinity of the leading edge (2) seem to indicate that the shock wave and boundary layer are merged for some distance before separating in a rarefied hypersonic flow. This flow model was assumed in (12) for the strong interaction region in originally deriving the induced pressure variation as being linear with X^{-1} , where X is the distance from the leading edge. The important interaction parameter was found to be $\chi = M_1^3/\sqrt{Re_x}$, where M_1 is the free stream Mach number and Re_x is the Reynolds number based on the distance X .

In (17) a careful investigation of the flow between the shock wave and the plate surface was conducted at a flow Mach number of 5.8. By heating the air to 1360 R, the induced pressures over a flat plate (18) were investigated at a Mach number of 9.6 with a maximum interaction parameter of about 10. A thin flat plate in a rarefied flow was investigated in the University of California Low Density Tunnel over a Mach number range of 2 to 5.8 and with a maximum interaction parameter of about 12 (1-3). Similar investigations for a flat plate have been conducted in helium tunnels (19, 20). The highest interaction parameter in helium was about 10. The only other available induced pressure results in air for a flat plate at high flow Mach numbers have been presented in (4) up to a flow Mach number of 19.6, and some results for a stagnation temperature of 7200 R have been obtained at a Mach number of 7.8.

The present investigation was conducted to extend the available knowledge regarding the slip phenomenon at the leading edge of a sharp flat plate, the interaction of the shock wave and boundary layer in rarefied hypersonic flows, and the effects of real gas upon the flow over a flat plate at high Mach numbers. The flow Mach number in the test section varied from 7.95 to 25.1 with reflected equilibrium stagnation temperatures of approximately 2300 and 7000 R. For these conditions, the calculated ambient mean free path varied from 0.0004 to 0.0859 in. With a leading edge thickness of approximately 0.001 in., the Knudsen number varied from 0.40 to 85.9 and the corresponding range of Reynolds number per inch was 37,900 to 326. The static pressure in the test section varied from 0.95 mm to 3.2 μ of mercury and the interaction parameter ranged from 2.5 to 756.

The leading edge slip phenomenon for a flat plate with a sharp leading edge has been investigated by measuring the surface pressure and obtaining schlieren photographs of the delay in the formation of the strong shock wave and boundary layer at the leading edge. The shock wave-boundary layer interaction results will be discussed for both nearly perfect gas and for a gas at a stagnation temperature of 7000 R where some real gas effects are present.

Description of Experiments

Hypersonic Shock Tunnel and Instrumentation

The tests were conducted in a hypersonic shock tunnel with a 4-in. ID and 103 ft long constant area driven tube and a reflected conical nozzle with an exit diameter of 24 in. A more detailed description of the facility is presented in (21). The shock reflects at the nozzle entrance and further increases the pressure and temperature. The throat diameter of the nozzle is changeable, allowing for a Mach number variation in the test section from 5.5 to 30. A scored aluminum diaphragm at the nozzle entrance permits the evacuation of the dump tank and nozzle to approximately 5 μ of mercury to facilitate the flow establishment and minimize starting shocks.

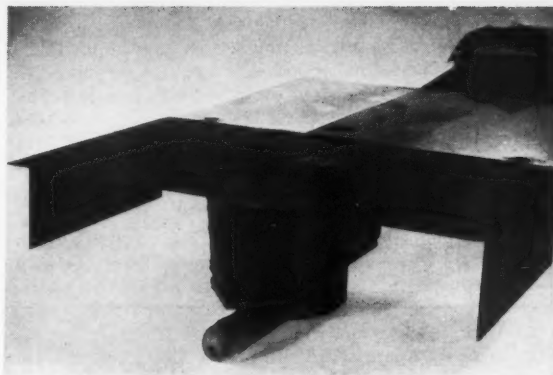


Fig. 1 Flat plate model

The model is mounted on a hollow 6-in.-diameter sting which is independently supported from the dump tank and floor by the use of Teflon bushings. The static pressure gages used in the model are barium-titanate piezoelectric gages which are small in size and high in output. They are very sensitive to position and mounting so they are dynamically calibrated at the end of the constant area driven tube by using the pressure rise across a weak incident shock wave.

Optical investigation of the model is possible through thick optically selected plate glass windows by means of a single pass schlieren system. A light source is provided by a 0.4 μ sec duration, 10,000 v capacitor discharge spark.

Model

The steel flat plate model tested in the hypersonic shock tunnel was 5 in. wide and 6 in. long. For high Mach numbers and low density conditions, side extension plates were also used to increase the overall width of the plate to 10 in. A 5- or 10-in. lucite extension was mounted flush with the trailing edge to minimize any base effects. Likewise, thin side plates, see Fig. 1, were attached to the plate to prevent any disturbance from the gage housing, mounted beneath the plate, affecting the flow over the top surface. The leading edge of the plate was observed under a microscope and the thickness was estimated to be 0.001 in.

There were five pressure orifices located along the surface of the plate at the following distances from the leading edge: 0.210, 0.772, 1.396, 2.795 and 4.008 in. The gages were all housed in the wedge mounted beneath the plate, Fig. 1. The impact pressure at the plane of the leading edge was measured by a Kistler gage in a cylindrical housing with a hemispherical nose. All gages were shock mounted and electrically insulated from the model and dynamically calibrated in the model before and after each series of tests.

Test Conditions

For a hypersonic shock tunnel with a combustion driver, it is possible to operate over a wide range of reflected equilibrium stagnation temperatures and pressures at the entrance to a conical nozzle (21). The equilibrium stagnation temperature and pressure behind the reflected shock wave are controlled by the strength of the incident shock wave and the initial temperature and pressure in the driven tube.

The pressure P_s behind the reflected shock wave at the entrance to the nozzle is measured with a standard Kistler quartz gage. The corresponding reflected stagnation temperature T_s is calculated using the known shock velocity at the end of the driven tube and the equilibrium thermodyna-

Table 1 Flow conditions

M_1	T_b R	P_1 μ of Hg	λ_1 $\times 10^3$ in.	λ_{max} $\times 10^3$ in.	λ_{min} $\times 10^3$ in.	Kn_1	Re_t
9.86	2480	67.2	0.397	3.05	0.739	0.397	37.9
15.2	2340	79.6	0.804	25.6	2.13	0.804	28.7
19.8	2310	15.9	1.89	129.0	4.82	1.89	15.8
25.1	2330	3.21	5.18	640.0	24.6	5.18	7.35
7.95	6750	945.0	3.21	21.5	0.536	3.21	3.75
12.2	7000	62.8	18.2	32.4	3.60	18.2	1.02
15.1	6900	14.9	41.8	137.0	9.78	41.8	0.546
18.5	6680	3.81	85.9	532.0	24.2	85.9	0.326

mic data for air (22-26). Other methods for determining the high temperature of gases are discussed in (27). For the present investigation, the reflected pressures were high enough to have nearly equilibrium conditions before the gas was expanded in the nozzle.

Recently a number of analytical papers (28-33) have considered the problem associated with the expansion of high temperature air in a hypersonic nozzle. The state of the air in the nozzle was determined by assuming recombination rates for the air constituents at elevated temperatures. These analytical results indicate some of the effects of stagnation pressure and temperature levels upon the expansion process. Unfortunately, there is no reliable data for the recombination rates of air at the elevated temperatures. To obtain some information for the air expansion process in a hypersonic nozzle, an extensive investigation of the static pressure and the impact pressure along the axis of the nozzle was conducted over a reflected equilibrium stagnation temperature range of 2100 to 11,000 R and with reflected pressures from 100 to 2000 psia. A detailed discussion of these results is presented in (34). The observed experimental pressure data were compared with the calculated equilibrium and "frozen" cases. It was found that the expansion of the air in the nozzle is essentially in equilibrium up to a reflected equilibrium stagnation temperature of about 8100 R for a reflected pressure greater than 500 psia (34). Close agreement for the present test conditions was obtained between the calculated values for equilibrium flow and the measured results from the static and impact pressures.

Leading Edge Slip Phenomenon in a Rarefied Hypersonic Flow

Flow Conditions at the Leading Edge Region

Knowing the temperature and density in the free stream, the mean free path λ defined by the kinetic theory of gases is given by

$$\mu = 0.499 \rho \bar{v} \lambda$$

where μ is the viscosity, ρ is the density, and \bar{v} is the average molecular velocity. At the low temperatures, the viscosity was assumed to be a linear function of the temperature and for the intermediate range of temperatures, Sutherland's equation (35) for viscosity was used. Fortunately, in the test section the highest free stream temperature was 744 R with a stagnation temperature of 6750 R at a Mach number of 7.95. Thus, the real gas effects upon the viscosity were negligible for the test section ambient temperatures, and the ratio of the specific heats γ was taken to be equal to 1.4 in determining the average molecular velocity. In Table 1 the values of the mean free path λ_1 in the test section are tabulated for the different flow conditions.

The mean free path λ_{max} at the leading edge of the flat plate was calculated by assuming that the pressure was equal

to the free stream pressure and the temperature was equal to the wall temperature of 540 R. This mean free path becomes the upper limit for the possible value on the plate surface. These values for the mean free path are presented in Table 1. For both stagnation conditions, the mean free path at the leading edge is much larger than that in the test section at high flow Mach numbers.

A short distance back of the leading edge, the surface pressure and the shock wave angles reach maximum values. By using this maximum pressure and the wall temperature of 540 R, the mean free path λ_{min} can be determined, see Table 1. Since the surface pressure downstream decreases from the maximum value, the mean free path of the air at the surface condition must increase with the distance from the leading edge for a constant wall temperature.

The Knudsen number, ratio of the free stream mean free path and the leading edge thickness, is presented in Table 1 for all the test conditions. For the high stagnation temperature conditions, the Knudsen numbers are very large for all the flow Mach numbers. At these large values of the Knudsen number, the flow at the leading edge is in the free molecular flow regime, where the collisions between the molecules are negligible compared to the collisions between the molecules and the plate surface. Even larger values for the Knudsen number are calculated using the mean free path λ_{max} at the plate leading edge with the gas temperature equal to the surface temperature.

The Reynolds numbers Re_t based upon the leading edge thickness of 0.001 in., and the ambient flow conditions are tabulated in Table 1. For all flow conditions, the Reynolds numbers are quite low and for the stagnation temperatures of about 7000 R, the Reynolds number is less than unity. At these low values and with the Knudsen number much greater than unity, the conventional boundary layer does not form at the leading edge region. The Reynolds number per inch indicates that at the high temperature conditions rarefied flow phenomena exist over the forward portion of the flat plate.

Surface Pressure Distribution Near the Leading Edge

The induced pressure ratios have been plotted in Figs. 2 and 4 for stagnation temperatures of 2300 and 7000 R as a function of the distance from the leading edge. In Fig. 2, the range of flow Mach number is from 9.86 to 25.1 with negligible real gas effects present because of the low stagnation temperature. All of the results for the flow Mach numbers less than 19.8 indicate the trend toward an asymptotic maximum pressure in the vicinity of the leading edge. The highest flow Mach number of 25.1 indicates the possible decrease in the induced pressure as the leading edge is approached. It is also apparent from this figure that the induced pressure on a flat plate, at the high flow Mach numbers, becomes appreciable over a large distance.

In (4), for a flat plate with a slightly thicker leading edge

and stagnation temperature of 7200 R, only the results at a flow Mach number of 7.8 were presented. Induced pressures for a flat plate have been obtained in the present investigation up to a flow Mach number of 18.5 with a stagnation temperature of 7000 R. The induced pressures tend to an asymptotic maximum value as the leading edge is approached. At the highest flow Mach number the pressure tends to drop off in the vicinity of the leading edge. For high flow Mach numbers and large Knudsen numbers, the flow at the leading edge

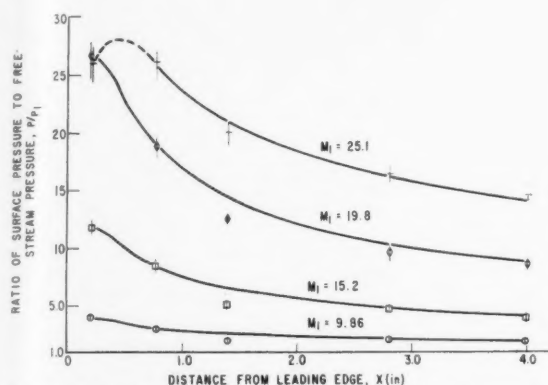


Fig. 2 Induced pressure on a flat plate vs. distance from the leading edge, $T_s = 2300$ R

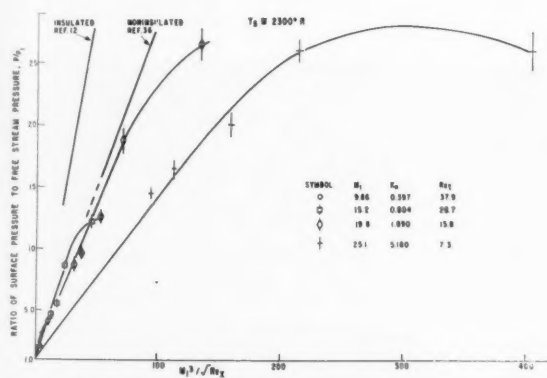


Fig. 3 Induced pressure on a flat plate vs. interaction parameter, $M_1^3/\sqrt{Re_x}$, $T_s = 2300$ R

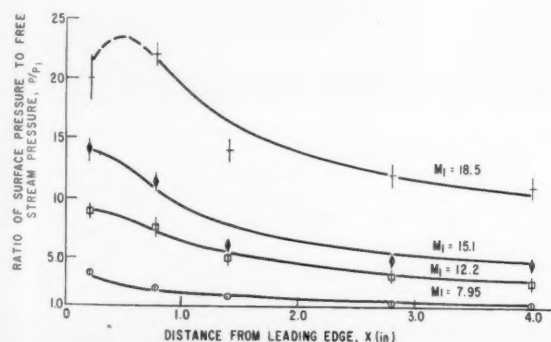


Fig. 4 Induced pressure on a flat plate vs. distance from the leading edge, $T_s = 7000$ R

exhibits the slip flow phenomenon, and the shock wave and the boundary layer is delayed in forming. Thus, the maximum induced pressure ratio occurs at a point downstream of the leading edge at high flow Mach numbers, Fig. 4.

The induced pressure ratios have been plotted in Figs. 3 and 5 as a function of the strong interaction parameter $M_1^3/\sqrt{Re_x}$ for stagnation temperatures of 2300 and 7000 R. For a stagnation temperature of 2300 R and a wall temperature of 540 R, the ratio of wall temperature to the recovery temperature is 0.28 for a laminar boundary layer. The schlieren photographs at these high flow Mach numbers have indicated that a laminar boundary layer exists on the plate surface. In Figs. 3 and 5 the theoretical predictions (12, 36) of the induced pressures for an insulated and noninsulated case are plotted. For the stagnation temperature of approximately 2300 R, the wall temperature ratios for all the flow Mach numbers were about the same. The induced pressure ratios all depart from the straight line variation in Fig. 3 at the leading edge region. For the highest flow Mach number of 25.1, the tapering off of the induced pressure near the leading edge is the most pronounced.

For the higher stagnation temperature of 7000 R, the departure from the straight line variation of the induced pressure with the interaction parameter $M_1^3/\sqrt{Re_x}$ is very pronounced at all of the flow Mach numbers, as indicated in Fig. 5. At the highest flow Mach number of 18.5 for this particular stagnation temperature condition, the leading edge Knudsen number and Reynolds number were 85.9 and 0.326, respectively. The interaction parameter for the first pressure orifice was 756. For this value, though the strong interaction theory would predict a very large induced pressure ratio, the actual observed maximum pressure ratio was approximately 28. This large decrease in the pressure from the strong interaction theory and the departure from the linear variation with $M_1^3/\sqrt{Re_x}$ indicates the existence of the slip flow phenomenon at the leading edge. The strong shock wave and the conventional boundary layer are delayed in forming over an appreciable distance for this particular rarefied hypersonic flow condition.

Similar induced pressure behavior in the vicinity of the leading edge has been reported for a flat plate in rarefied flow at a flow Mach number range of 3.5 to 5.8 (1-3). At a flow Mach number of 5.8, the surface pressure was measured at the highest interaction parameter of 8 and the induced pressure ratio corresponding to this location was 4.8. For this flow condition the observed induced pressure began to depart from a straight line for an interaction parameter of approximately 7.8. The value of the interaction parameter at which the induced pressure departs from the linear variation with $X^{-1/2}$ increases drastically with the increasing flow Mach number, as indicated in Figs. 3 and 5.

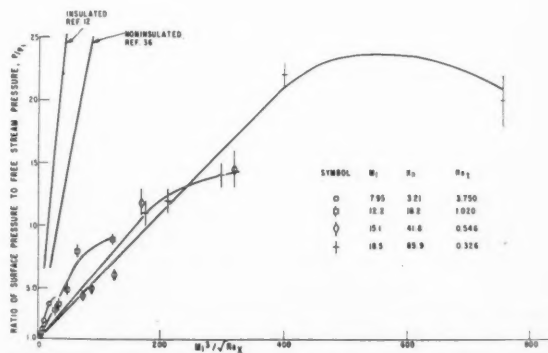


Fig. 5 Induced pressure on a flat plate vs. interaction parameter, $M_1^3/\sqrt{Re_x}$, $T_s = 7000$ R

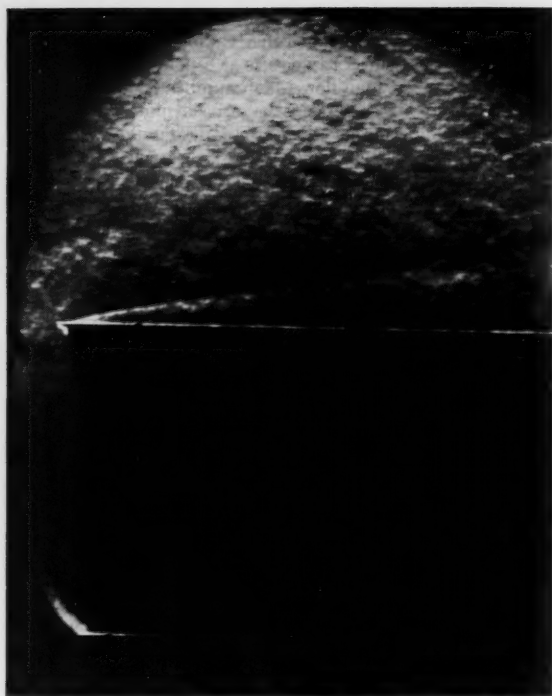


Fig. 6 Schlieren photograph of flow over a flat plate at Mach number, 19.8, $T_s = 2300$ R

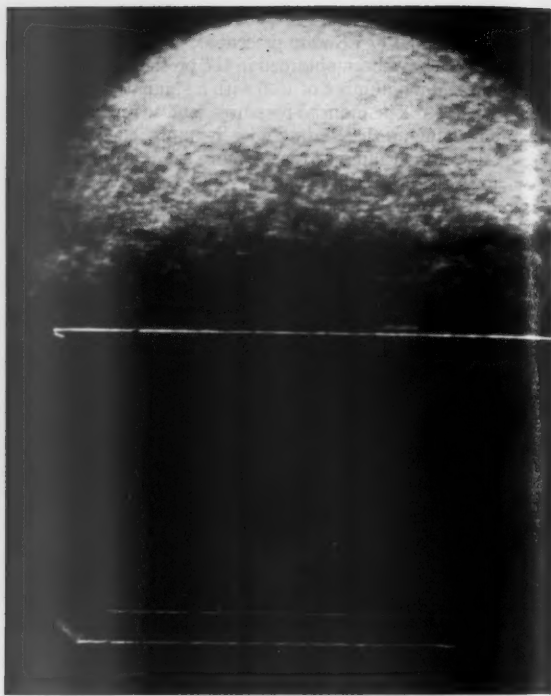


Fig. 7 Schlieren photograph of flow over a flat plate at Mach number 25.1, $T_s = 2300$ R

Schlieren Photographs

All the schlieren photographs, Figs. 6-8, were taken with a single-pass system equipped with a 0.4 microsec spark source. For the majority of the present tests, the span of the model was 10 in. to obtain the maximum optical length through the shock wave and the boundary layer. In these schlieren photographs, the outer edge of the boundary layer is visible because of the large density gradient (17) existing at the outer edge for a hypersonic laminar boundary layer.

For a stagnation temperature of 2300 R, it was possible to obtain schlieren photographs of the flow over the plate from Mach number 9.86 to 25.1. At the highest flow Mach number, the static pressure in the test section was only 3.2μ of mercury and the mean free path was 0.00518 in. The maximum shock wave angle did not occur at the leading edge, but was located approximately 0.5 in. back of the tip, Fig. 7. The shock wave and the boundary layer are more clearly seen at a lower flow Mach number of 19.8 because of the higher density in the test section, Fig. 6. For this flow condition, the Knudsen number was 1.89 with an ambient mean free path of 0.00189 in. The maximum shock wave angle was delayed in forming but was located closer to the leading edge than for the higher flow Mach number of 25.1. After the location of the maximum shock wave angle, the shock wave and boundary layer are merged before separating, as indicated in Fig. 6. At the higher flow Mach numbers, the maximum shock wave angles are located downstream of the leading edge instead of being at the leading edge $X = 0$ as predicted by the classical boundary layer theory with no slip at the surface. The length of the shock wave and boundary layer merged region increased with the flow Mach number, as indicated by the schlieren photographs. Also, the location of the reverse curvature in the shock wave and boundary layer moved downstream from the leading edge with the increasing flow Mach number and lower ambient density.

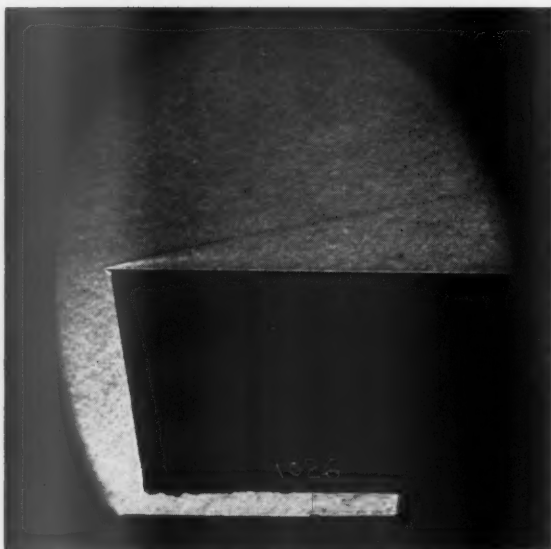


Fig. 8 Schlieren photograph of flow over a flat plate at Mach number 7.95, $T_s = 7000$ R

For higher stagnation temperatures, it was possible to obtain schlieren photographs, Fig. 8, only at the lower Mach numbers due to the extremely rarefied flow conditions. Fig. 8 is for a flow Mach number of 7.95 and a stagnation temperature of 7000 R. The stagnation temperature of 7000 R corresponds approximately to the temperature encountered

during re-entry by intermediate range ballistic missiles. The shock wave and the boundary layer are delayed in forming, Fig. 8, and the maximum shock angle is located downstream of the leading edge. For a given Reynolds number, as the mean free path ahead of the model becomes larger and the temperature of the gas higher, the delay in the maximum shock angle becomes larger and the boundary layer thicker even with a cooled wall condition.

By measuring the shock wave angle from the schlieren photograph, the pressure ratio across the shock can be calculated by using the oblique shock wave relationship for perfect gas as well as for real gas with thermodynamic equilibrium (37). The pressure ratio across the shock wave increased from the leading edge to a peak value corresponding to the maximum shock wave angle. After this maximum value, the pressure ratio continuously decreased as the shock wave angle became smaller. For a flow Mach number of 25.1, the pressure ratio near the leading edge was approximately one-half the maximum pressure ratio value downstream from the leading edge. The results for the higher stagnation temperature of 7000 R have indicated a lower pressure ratio at the leading edge compared to the peak pressure ratio. The increase in the pressure ratio behind the shock wave to a maximum pressure ratio could be caused by the slip flow existing at the leading edge region (1,3,4).

Leading Edge Slip-Flow Effects

For high flow Mach numbers in the test section and at high stagnation temperature conditions, the static pressure in the test section is in the micron range and the mean free path is quite large, see Table 1. For these conditions, the flow over the flat plate is in the low density hypersonic flow region with some real gas effects in the boundary layer for the 7000 R stagnation temperature.

When the Knudsen number is equal to or greater than 1.0, the flow over a flat plate can be considered to be the case for a sharp leading edge. When the Knudsen number is much greater than 1, the flow at the leading edge can be considered as being in the free molecular regime. For this condition the collisions between the molecules can be neglected compared to the collisions of the molecules with the leading edge and the plate surface. For the Knudsen number equal to or greater than unity in hypersonic flow, the possibility of forming a detached shock wave and a conventional boundary layer is remote.

Besides the Knudsen number, the flow Mach number is also an important parameter for the existence of slip flow phenomenon at the leading edge. By assuming the hypersonic flow to be Maxwellian (38) and $Kn \geq 1.0$, the molecules will travel a distance equal to the product of the Mach number and the ambient mean free path in the time of one collision for the free stream condition. It will require time for N collisions to bring the velocity at the surface to zero and as a first approximation the slip distance can be given by

$$\zeta = M\lambda_1 N$$

as derived in (5). Thus, for flow Mach numbers, $M \gg 1.0$, the particles travel an appreciable distance along the plate even for the time required for one collision between the molecules in the free stream.

The induced pressure on a flat plate due to the interaction of the shock wave and the viscous layer increases as the square of the free stream Mach number for a given flow deflection. This is apparent from the oblique shock wave equation

$$P_2/P_1 = \frac{2\gamma}{\gamma+1} M_1^2 \sin^2 \beta - \frac{\gamma-1}{\gamma+1}$$

With real gas effects after the oblique shock wave, this equation can still be applied with the correct ratio of the specific heats for the assumption of equilibrium flow behind the shock wave. If the leading edge is blunt compared to the mean

free path $Kn \ll 1.0$, a detached shock wave can form ahead of the plate and the pressure after the normal shock wave on the axis becomes very large at high flow Mach numbers. Hence, at high flow Mach numbers the difference in the maximum induced pressure for a sharp and blunt flat plate can be very large.

When the Reynolds number, based upon the leading edge thickness and the free stream condition, is of the order unity, and the Knudsen number is much greater than unity, the boundary layer at the leading edge will not form in the conventional manner (2, 39). Hence, for this type of flow condition it is not possible to apply the conventional boundary layer results in the vicinity of the leading edge. For the present investigation at high stagnation temperatures and flow Mach numbers, the leading edge Reynolds number was less than 1.0 and the Knudsen number was much greater than unity. For these conditions the slip flow seems to exist at the leading edge region and the shock wave and boundary layer are merged after the slip region, Figs. 6-8. For the lower stagnation temperature case, the leading edge Reynolds number is about 10 at the high Mach numbers but the Knudsen number is greater than unity so that slip flow phenomenon can exist, Figs. 6 and 7.

For large Knudsen numbers, low leading edge Reynolds numbers and high flow velocities, the flow at the leading edge is nearly free molecular, followed by a rarefied flow and finally, at the end of the slip distance, the continuum flow with velocity at the wall equal to zero, as indicated in Fig. 9. In the slip flow region, the velocity at the wall decreases from nearly the free stream value to zero at the end of the slip distance (39). The shock wave slope increases from the leading edge to a maximum value at the end of the slip region. Corresponding to this shock wave behavior, the pressure on the surface must also increase from a finite value at the leading edge to a maximum value slightly downstream of the leading edge before decreasing as $X^{-1/2}$. The shock wave and boundary layer are both merged and delayed in forming in the slip region. These results for the slip flow region, based upon the kinetic theory of gases, have been discussed in (5).

Shock Wave-Boundary Layer Interaction in High Temperature Rarefied Hypersonic Flow

Induced Pressure Distribution for Reflected Stagnation Temperature of 2300 R

The induced pressure results for the stagnation temperature of approximately 2300 R and flow Mach number of 9.86 to 25.1 are presented in Figs. 2 and 3. For these investigations, the flat plate surface temperature was close to room temperature throughout the test duration. All the schlieren photographs of the flow have indicated a laminar boundary layer to exist on the plate so that a recovery factor of 0.85 was used to calculate the recovery temperature at the surface. The ratio of the wall temperature to the recovery temperature is approximately 0.28 for this Mach number range. For these conditions the cooling effects of the plate surface upon the induced pressures caused by the interaction of the shock wave and the boundary layer must be considered. In (36) the effects of both heating and cooling upon the induced pressure, skin friction and heat transfer have been derived for a flat plate with a sharp leading edge. At this particular stagnation temperature the real gas effects upon the induced pressure are small compared to the wall cooling effects.

The vertical bars of the test data give the variation in the observed induced pressure ratios during the test. At higher flow Mach numbers, where the ambient pressure in the test section is low, see Table 1, the variation of the test data during a given run becomes greater because of the large gain required to amplify the output of the piezoelectric gages. With additional improvement in the shock isolation of the

model and the gage, it will be possible to improve the signal to noise ratio at these conditions. For the lower Mach numbers, the Reynolds number per inch was high so that the amount of the induced pressure variation for a given test became small for both stagnation temperatures of 2300 and 7000 R.

The asymptotic values of the induced pressure ratio, as X becomes larger, approach unity for all the flow Mach numbers. These results also indicate that the lucite extension at the end of the flat plate and the fences at the sides of the plate, as shown in Fig. 1, were effective in reducing the external disturbances on the boundary layer.

For better representation of the induced pressure effects in the physical plane, the measured pressure ratios are plotted in Fig. 2 as a function of the distance from the leading edge. It is apparent from this figure that the interaction of the shock wave with the boundary layer raises the pressure level on the flat plate. For flow Mach numbers up to 19.6, the first pressure orifice was located back of the slip flow region so that the induced pressure continuously decreased over the plate surface with increasing X . But, for the flow Mach number of 25.1, the first pressure orifice, located at $X = 0.210$ in., was in the leading edge slip region and, hence, the pressure at the second orifice, $X = 0.772$ in., was about the same as the first orifice location. At this extremely high flow Mach number, the maximum pressure ratio was about 28 and even at a distance of 4 in. back from the leading edge the pressure was 14 times the free stream value ahead of the plate. Due to this viscous interaction phenomena, the skin friction and heat transfer is much higher than the values determined by neglecting the induced pressure effects (9-16, 36). The force and the stability for a lifting surface at high flight Mach numbers is greatly affected by the interaction of the shock wave and viscous layer. The flow condition at the leading edge seems to determine the pressure level at the very high Mach numbers.

In Fig. 3 the induced pressure ratio has been plotted as a function of the interaction parameter $\chi = M_1^3/\sqrt{Re_x}$. For the flow Mach numbers that were investigated, the interaction parameter varied from about 2.5 to 404. For nearly perfect gas conditions, $\gamma = 1.4$, the induced surface pressures have been measured by Kendall (17) at Mach 5.8 and $\chi = 3.1$, Betram (18) at Mach number 9.6 and $\chi = 8.5$, Schaff et al. (1-3) at Mach number 5.8 and $\chi = 10$. Some plate measurements had been made in helium, $\gamma = 1.67$, by Bogdonoff and Hammitt (19) at Mach number 13.6 and $\chi = 6$, and Erickson (20) at Mach number 17.3 and $\chi = 10$. In our previous paper (4) the highest flow Mach number was 19.6 with the maximum χ value of 140. The amount of available data in air for flow Mach numbers above 15 is still rather limited.

For each flow Mach number, the induced pressure ratio approaches an asymptotic value of different magnitude, indicating the existence of the slip phenomenon near the leading edge. These slip effects extend over an appreciable distance at a flow Mach number of 25.1, Fig. 3. Not only has the maximum pressure been reduced, but the slope of the induced pressure as a function of the interaction parameter has been drastically lowered. Laurmann (3) has made a first approximation calculation of the induced pressure due to the slip at the surface and the results indicate the leveling off of the pressure as the leading edge is approached. As the flow Mach number is increased, the leading edge slip region is extended (5) since one of the parameters for the slip is the product of Mach number and mean free path. The slope of the straight portion of the induced pressure ratio in Fig. 3 decreases as the Mach number increases. The theoretical slope predicted for the insulated case (12) is also plotted. At a flow Mach number of 9.86, the theoretical slope for an insulated condition is reasonably close to the observed experimental results for a noninsulated case. The results obtained in the low density wind tunnel (1-3) for induced pressures on

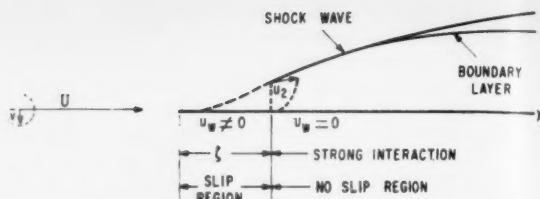


Fig. 9 Hypersonic flow near the leading edge of a flat plate

a flat plate agree reasonably well with the present $M = 9.86$ data.

The effects of noninsulated conditions upon the induced pressure, skin friction and heat transfer have been determined analytically (36) by assuming a perfect gas. For a stagnation temperature of 2300 R, the gas can be assumed to be close to a perfect gas without appreciable error. In Fig. 3 the theoretical slope for the induced pressure ratio as a function of the interaction parameter is plotted for the actual temperature ratio for the plate surface. At a flow Mach number of 15.2, the slope is slightly greater than the analytical value for the noninsulated condition and at a flow Mach number of 19.8, the slope is slightly less than the theoretical curve. For the flow Mach number of 25.1 with a Knudsen number of 5.2, the slope of the induced pressure ratio is much less than the theoretical slope. Most of this large decrease in the slope could be caused by the appreciable region of slip flow phenomenon at the leading edge due to the high flow Mach number, the large mean free path ahead of the flat plate, and the low Reynolds number based upon the leading edge thickness. The schlieren photograph of this flow, Fig. 7, has indicated the appreciable delay in the formation of the maximum shock wave angle and the merged region of the shock wave and boundary layer. These results seem to indicate that the leading edge slip flow effects can be appreciably greater than the surface cooling effects at high flow Mach numbers with rarefied conditions. Further experimental and analytical investigations must be conducted to understand the effects of surface cooling and leading edge slip-flow upon the induced pressure for a flat plate with a sharp leading edge.

Real Gas Effects on Shock Wave-Boundary Layer Interaction for Reflected Stagnation Temperature of 7000 R

The real gas effects upon the interaction of the shock wave and the boundary layer were investigated for a reflected stagnation temperature of approximately 7000 R and a stagnation pressure of about 500 psia. Previous investigations of the nozzle flow for these stagnation conditions with impact and static pressure probes have indicated that the flow expansion was very close to the equilibrium state (34). To investigate the flow in the test section for the above stagnation temperature and area ratio of 576, an adjustable wedge with pressure gages was tested. The shock wave angle and the surface pressure were compared with the oblique shock wave calculations for equilibrium flow with good agreement (37). All of these investigations have indicated that for the above stagnation temperature and pressure, the flow in the nozzle is very close to the equilibrium case.

At this elevated stagnation temperature, the Reynolds number per inch is quite low and the mean free path in the test section for different area ratios is appreciable (Table 1). For a leading edge thickness of approximately 0.001 in., the leading edge Reynolds numbers based upon the free stream condition are all less than unity except for the lowest Mach number of 7.95. The Knudsen numbers are all greater than one and the highest value was 85.9. For these large Knudsen

numbers and the low Reynolds numbers, the flat plate can be considered sharp with no possibility of forming a detached shock wave at the leading edge. The plate surface temperature for the duration of the flow was close to room temperature. By assuming the boundary layer to be laminar and a recovery factor of 0.85, the ratio of the wall temperature to the recovery temperature is 0.0872.

For this relatively high temperature and very cool surface compared to the free stream stagnation temperature, the induced pressure caused by the interaction of the leading edge shock wave and the boundary layer is present, as indicated by Figs. 4 and 5. For all the flow Mach numbers above 7.95, the pressure near the leading edge departed noticeably from the linear variation with $X^{-1/2}$. At the lower stagnation temperature, only the highest flow Mach number of 25.1 indicated the large leading edge slip phenomenon. The maximum induced pressure ratios seem to again increase approximately as the square of the Mach number. After the leading edge slip region, the induced pressure in the strong interaction region varied as $X^{-1/2}$.

In Fig. 4 the induced pressure ratio is plotted as a function of the distance from the leading edge X for the different hypersonic flow Mach numbers. The results indicated that even at this relatively high stagnation temperature the induced pressure over the flat plate is appreciable. Due to the large Knudsen number and low Reynolds number at the flow Mach number of 18.5, the pressure at the first orifice location is slightly less than that at the second orifice. This result indicates that the slip flow phenomenon extended over an appreciable distance back of the leading edge. At even higher flow Mach numbers and rarefied flow conditions, the leading edge slip phenomenon may extend over a large percentage of the plate surface. This high induced pressure over the plate surface will increase the skin friction, heat transfer, and forces acting on a lifting surface compared to the case without the shock wave-boundary layer interaction. The reflected stagnation temperature of 7000 R corresponds approximately to the temperature after the normal shock wave encountered by a vehicle traveling at a flight Mach number of approximately 16 at 100,000 ft altitude. Thus, in the design of the hypersonic vehicle with lifting surfaces, the leading edge slip phenomenon and the induced pressure due to the interaction of the shock wave and the boundary layer must be considered.

The induced pressure ratios are plotted as a function of the strong interaction parameter $M_1^3/\sqrt{Re_x}$ for the different flow Mach numbers in Fig. 5. For the strong interaction region with surface cooling, the induced pressure ratio is a linear function of the above parameter. For all the flow Mach numbers that were investigated, the induced pressure ratio departs very noticeably from the linear variation as the interaction parameter increases. This departure from the linear variation is probably caused by the slip flow phenomenon (1,3-5).

The induced pressure on the plate will increase from the leading edge to a maximum value at the end of the slip flow region. As the flow Mach number in rarefied gas becomes very large, the slip flow effects can extend over an appreciable distance from the leading edge. The theoretical value for the slope of the induced pressure was derived on the basis of a cooled surface, perfect gas and no slip at the surface (36). By comparing this value with the actual experimental results, (Fig. 5), it is apparent that the slip flow effects are very large at high flow Mach numbers, compared to the surface cooling effects. At the lower flow Mach number of 7.95, the observed experimental data is closer to the theoretical curve for the noninsulated case. The slip extent is governed by the product of the Mach number and the mean free path so that at the lower flow Mach number the slip flow effects will be smaller than at very high Mach numbers.

These results for the flat plate with a sharp leading edge at hypersonic Mach numbers in rarefied, high temperature

air indicate that the leading edge slip phenomenon and the interaction of the shock wave and boundary layer have a definite influence upon the flow. The free stream Mach number plays an important role on the extent of the leading edge slip region and the level of the maximum induced pressure. The mean free path in the free stream is an important factor for determining sharpness of the leading edge, $Kn \geq 1.0$, and the extent of the slip region (5), $\zeta = M\lambda N$. A cool surface will tend to decrease the boundary layer thickness and lower the induced pressure on the surface. Since the experimental results for a flat plate at high flow Mach numbers and a rarefied gas do not agree with the existing strong interaction theories (10-16), which neglect the leading edge slip phenomenon, further investigations must be conducted to develop a theory which includes these physical parameters.

Conclusion

The flow over a flat plate with a sharp leading edge has been investigated over a wide range of test conditions in the hypersonic shock tunnel. The flow Mach number varied from 9.86 to 25.1 for a nearly perfect gas condition with a stagnation temperature of approximately 2300 R. To investigate the real gas effects upon the flow over a flat plate, a stagnation temperature of approximately 7000 R was utilized and the corresponding flow Mach number varied from 7.95 to 18.5.

Schlieren photographs and induced pressure measurements have indicated the presence of slip flow at the leading edge of the plate. The observed results seem to indicate that the slip distance is dependent upon the free stream Mach number, the mean free path, and the number of collisions necessary to decrease the velocity at the wall to zero. The schlieren photographs indicate the delay in formation of the strong shock wave and boundary layer near the leading edge. The surface pressures also deviated from the strong interaction theory in the vicinity of the leading edge.

In the strong shock wave-boundary layer interaction region, the viscous layer and the shock wave were merged and there was no apparent inviscid region between them. After the maximum induced pressure, the surface pressures varied inversely as the square root of the distance from the leading edge, as originally predicted by Nagamatsu and Li. The deviation from this theoretical result may be caused by the occurrence of slip flow at the leading edge region. These results were observed for both the nearly perfect gas condition and the higher stagnation temperature with real gas effects. The maximum induced pressure on the surface of the plate was 28 times the free stream pressure and even at the lowest flow Mach number of 7.95, the induced pressure was four times the ambient pressure.

The strong interaction parameter χ varied from 2.5 to 756 and the induced pressure results are compared with the theoretical predictions of Li and Nagamatsu for a noninsulated plate. While there seems to be a reasonable agreement for the low stagnation temperature condition, the experimental results deviate appreciably from the theoretical curve at the higher stagnation temperature as the Mach number and Knudsen number increase. This deviation may be caused primarily by the slip flow phenomenon instead of the wall cooling effects, but more investigations must be conducted before any definite conclusions can be reached.

The pressure ratios determined from the schlieren photographs by measuring the shock wave angles have indicated the increase in the pressure from the leading edge to a maximum value slightly downstream; thereafter the pressure decreases. When the shock wave and the boundary layer were merged, the pressure at the shock wave and the plate surface were about equal. After the shock wave separated from the boundary layer, the pressure behind the shock wave was greater than the surface pressure at a given distance from the leading edge.

High induced pressures, caused by the interaction of the shock wave and the boundary layer, increase the skin friction, heat transfer, and force on a lifting surface as compared to the values calculated by the compressible laminar boundary layer theory with no pressure gradient. These effects will be appreciable during the re-entry of glide-type vehicles at satellite velocities.

Acknowledgment

The enthusiasm and support of A. J. Nerad contributed greatly to the attainment of the results presented in this paper. T. C. Flaherty assisted in the operation and reduction of data. L. A. Osburg and K. H. Cary assisted with the mechanical design and instrumentation.

References

- 1 Schaaf, S. A., Hurlbut, F. C., Talbot, L. and Aroesty, J., "Viscous Interaction Experiments at Low Reynolds Numbers," *ARS JOURNAL*, vol. 29, no. 7, July 1959, pp. 527-528.
- 2 Laurmann, J. A., "The Free Molecule Probe and Its Use for the Study of Leading Edge Flows," *Phys. Fluids*, vol. 1, no. 6, 1958, pp. 469-477.
- 3 Laurmann, J. A., "The Effect of Slip on Induced Pressures," *J. Aero/Space Sci.*, vol. 26, no. 1, 1959, pp. 53-54.
- 4 Nagamatsu, H. T. and Sheer, R. E., Jr., "Hypersonic Shock Wave-Boundary Layer Interaction and Leading Edge Slip," *ARS JOURNAL*, vol. 30, no. 5, May 1960, pp. 454-462.
- 5 Nagamatsu, H. T. and Li, T. Y., "On Hypersonic Flow Near the Leading Edge of a Flat Plate," *Phys. Fluids*, vol. 3, 1960, p. 140.
- 6 Hammitt, A. G., "The Hypersonic Viscous Effects on a Flat Plate with Finite Leading Edge," *J. Fluid Mech.*, vol. 5, no. 2, 1959, pp. 242-256.
- 7 Betram, M. H., "Viscous and Leading-Edge Thickness Effects on the Pressures on the Surface of a Flat Plate in Hypersonic Flow," *J. Aero. Sci.*, vol. 21, no. 6, 1954, pp. 430-431.
- 8 Lees, L. and Kubota, T., "Inviscid Hypersonic Flow Over Blunt-Nosed Slender Bodies," *J. Aero. Sci.*, vol. 24, no. 3, 1957, pp. 195-202.
- 9 Cheng, H. K., Hall, J. G., Golian, T. C. and Hertzberg, A., "Boundary Layer Displacement and Leading Edge Bluntness Effects in High Temperature Hypersonic Flow," *J. Aero/Space Sci.*, vol. 28, no. 5, 1961, pp. 353-381.
- 10 Shen, S. F., "An Estimate of Viscosity Effects on Hypersonic Flow Over an Insulated Wedge," *J. Math. & Phys.*, vol. 31, no. 3, 1952, pp. 192-205.
- 11 Betram, M. H., "An Approximate Method for Determining the Displacement Effects and Viscous Drag of Laminar Boundary Layers in Two-Dimensional Hypersonic Flow," NACA TN 2773, 1952.
- 12 Li, T. Y. and Nagamatsu, H. T., "Shock Wave Effects on the Laminar Skin Friction of an Insulated Flat Plate at Hypersonic Speeds," *J. Aero. Sci.*, vol. 20, no. 5, 1953, pp. 345-355.
- 13 Lees, L., "On the Boundary Layer Equations in Hypersonic Flow and Their Approximate Solutions," *J. Aero. Sci.*, vol. 20, no. 2, 1953, pp. 143-145.
- 14 Stewartson, K., "On the Motion of a Flat Plate at High Speeds in a Viscous Compressible Fluid—II Steady Motion," *J. Aero. Sci.*, vol. 22, no. 5, 1955, pp. 303-309.
- 15 Pai, S. I., "A Note on Hypersonic Flow Over a Flat Plate," *J. Aero/Space Sci.*, vol. 27, no. 7, 1953, pp. 502-503.
- 16 Oguchi, H., "First Order Approach to a Strong Interaction Problem in Hypersonic Flow Over an Insulated Flat Plate," Univ. of Tokyo, Aero. Res. Inst. Rep. no. 330, June 1958.
- 17 Kendall, J. M., Jr., "An Experimental Investigation of Leading-Edge Shock Wave-Boundary Layer Interaction at Mach 5.8," *J. Aero. Sci.*, vol. 24, no. 1, 1957, pp. 47-56.
- 18 Betram, M. H., "Boundary-Layer Displacement Effects in Air at Mach Numbers of 6.8 and 9.6," NACA TN 4133, 1958.
- 19 Bogdonoff, S. M. and Hammitt, A. G., "Fluid Dynamic Effects at Speeds from $M = 11$ to 15," *J. Aero. Sci.*, vol. 23, no. 2, 1956, pp. 108-116.
- 20 Erickson, W. D., NACA TN 4113, 1957.
- 21 Nagamatsu, H. T., Geiger, R. E. and Sheer, R. E., Jr., "Hypersonic Shock Tunnel," *ARS JOURNAL*, vol. 29, no. 5, May 1959, pp. 332-340.
- 22 Gilmore, F. R., "Equilibrium Composition and Thermodynamic Properties of Air to 24,000°K," Rand Rep. RM-1543, 1953.
- 23 Hilsenrath, J. and Beckett, C. W., "Tables of Thermodynamic Properties of Argon-Free Air to 15,000°K," Arnold Engng. Dev. Ctr., TN 56-12, 1956.
- 24 Treanor, C. E. and Logan, J. G., Jr., "Tables of Thermodynamic Properties of Air from 3000°K to 10,000°K," Cornell Aero. Lab. Rep. AD-1052-A-2, 1956.
- 25 Hirschfelder, J. D. and Curtiss, C. F., "Thermodynamic Properties of Air," Univ. of Wisconsin, CM-518; Navy Bu. Ord. Contract No. 9938, 1948.
- 26 Feldman, S., "Hypersonic Gas Dynamic Charts for Equilibrium Air," AVCO Res. Lab. Rep., Jan. 1957.
- 27 Kostkowski, H. J., *Proceedings of an International Symposium on High Temperature Technology*, McGraw-Hill Book Co., Inc., N. Y., 1960, pp. 33-44.
- 28 Logan, J. G., "Relaxation Phenomena in Hypersonic Aerodynamics," IAS preprint No. 728, 1957.
- 29 Bray, K. N. C., "Atomic Recombination in a Hypersonic Wind Tunnel Nozzle," *J. Fluid Mech.*, vol. 6, 1959, pp. 1-32.
- 30 Li, T. Y., "Non-Equilibrium Flow in Gas Dynamics," Rensselaer Polytechnic Inst., Rep. TR-AE-5901, 1959.
- 31 Hall, J. G. and Russo, A. L., "Studies of Chemical Non-Equilibrium in Hypersonic Nozzle Flow," Cornell Aero. Lab. Rep. AD-1118-A-6, Nov. 1959.
- 32 Bray, K. N. C. and Appleton, J. P., "The Choice of an Optimum Set of Measurements to Study Atomic Recombination in Nozzles," Univ. of Southampton, Rep. 120, 1959.
- 33 Lighthill, M. J., "Dynamics of a Dissociating Gas, Part I, Equilibrium Flow," *J. Fluid Mech.*, vol. 2, 1957, pp. 1-32.
- 34 Nagamatsu, H. T., Workman, J. B. and Sheer, R. E., Jr., "Hypersonic Nozzle Expansion with Air Atom Recombination Present," *J. Aero/Space Sci.* (in press).
- 35 "Equations, Tables and Charts for Compressible Flow," NACA TR 1135, 1953.
- 36 Li, T. Y. and Nagamatsu, H. T., "Hypersonic Viscous Flow on Non-Insulated Flat Plate," Proc. Fourth Midwestern Conf. on Fluid Mech., Purdue Engng. Research Series No. 128, pp. 273-287.
- 37 Nagamatsu, H. T., Workman, J. B. and Sheer, R. E., Jr., "Oblique Shock Relations for Air at Mach Number 7.8 and 7200°R Stagnation Temperature," *ARS JOURNAL*, vol. 30, no. 7, July 1960, pp. 619-623.
- 38 Kennard, E. H., *Kinetic Theory of Gases*, McGraw-Hill Book Co., Inc., N. Y., 1938.
- 39 Charwat, A. F., "Molecular Flow Study of the Hypersonic Sharp Leading Edge Interactions," Rand Corp. Rep. P-2113, Aug. 1960.

Analysis of Convective Heat Transfer in Rocket Nozzles

ERNEST MAYER¹

National Engineering Science Co.
Pasadena, Calif.

This paper presents simplified analytic formulas for the estimation of convective heat transfer to the cooled walls of rocket nozzles of novel configuration, such as spike nozzles, E-D and other external-expansion nozzles. In contrast with conventional nozzles, the exhaust gas flow along the walls of novel configuration cannot, in general, be referred to effective local stream diameters, and therefore these nozzles cannot be treated by application of conventional heat transfer formulas corresponding to fully developed turbulent flow in tubes and annuli. The heat transfer formulas derived in this paper are based on an approximate solution of the energy integral equation for boundary layers by Ambrok's method. The simplifying assumptions in the analysis correspond to conditions representative of rocket nozzle flow. Use is made of semi-empirical data associated with the Blasius flat plate-heat transfer coefficient, with modifications for variable fluid properties based on Eckert's reference temperature method. The calculation procedure requires as input data the cooled-wall temperature and the local free stream data just outside of the boundary layer, as determined by the inviscid analysis of the gas flow field. An illustrative example of the procedure is given for a representative Rocketdyne E-D nozzle configuration. Comparison of the predictions based on the boundary layer treatment in application to deLaval nozzles with those obtainable from conventional formulas applicable to this case shows reasonable agreement. However, a dependence of nozzle heat transfer on chamber configuration, not predicted by conventional formulas, is obtained by the present analysis, and is supported by reference to recent experimental data.

THE OBJECTIVE of this paper is to present formulas for the estimation of convective heat transfer coefficients in rocket nozzles, particularly in nozzles of novel configuration such as spike, E-D, and other external-expansion nozzles (1,2).² In contrast with the gas flow in the conventional deLaval nozzle, the exhaust gas flow in external-expansion nozzles cannot, in general, be referred to any well defined local stream diameters. Therefore, these nozzles cannot be treated by use of the heat transfer correlation based on fully developed turbulent flow in tubes (3), which has proved adequate in application to deLaval nozzles (4).

In this paper, rocket nozzle heat transfer formulas are derived by approximate solution of energy-integral equations for boundary layers along two-dimensional curved surfaces and surfaces of revolution. The energy integral considered corresponds to the von Kármán integral relation for momentum transfer (5). A rather general method of solving the energy-integral equation has been given by Ambrok (6). The solution given here is obtained by Ambrok's procedure in application to the special case of isentropic gas flow with boundary surface at constant temperature; i.e., conditions representative of rocket nozzle flow with cooled walls.

The Energy-Integral Equation

In the derivations which follow, the energy-integral equations of the boundary layer are expressed with the aid of the control volume sketch in Fig. 1, employing the notation of [(5), chapters 26, 27]. It is assumed for the present, that the fluid properties, specific heat c_p , dynamic viscosity μ and Prandtl number Pr are constant. The constant fluid prop-

erty values initially assumed will be designated with the subscript ∞ referring to free stream values. Allowance for variable fluid properties is made subsequently by use of Eckert's reference temperature method (7).

Consider the expression for the steady state heat flux from a free stream at constant stagnation temperature $T_{0,\infty}$ in longitudinal flow along a flat plate (see Fig. 1)

$$q_w = c_{p,\infty} T_{0,\infty} \frac{d}{dx} \int_0^\delta \rho u \left(1 - \frac{T_0}{T_{0,\infty}} \right) dy \quad [1]$$

where the integrand contains the boundary layer variables³

- y = coordinate normal to plate
- $\delta(x)$ = boundary layer thickness
- $\rho(x, y)$ = fluid density
- $u(x, y)$ = fluid velocity
- $T_0(x, y)$ = stagnation temperature

At the edge of the boundary layer $y = \delta$, the fluid variables attain the values corresponding to specified free stream data $\rho_\infty, U_\infty, T_{0,\infty}$. For boundary layers along curved plates, Eq. 1 becomes, to a high degree of approximation

$$q_w = c_{p,\infty} T_{0,\infty} \frac{d}{ds} \int_0^\infty \rho u \left(1 - \frac{T_0}{T_{0,\infty}} \right) dy \quad [2]$$

where y is the normal to the curvilinear coordinate s . The corresponding expression for surfaces of rotational symmetry is

$$q_w = \frac{c_{p,\infty} T_{0,\infty}}{r} \frac{d}{ds} r \int_0^\infty \rho u \left(1 - \frac{T_0}{T_{0,\infty}} \right) dy \quad [3]$$

where the meridional coordinate s and the radius of revolution

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ Associate Director of Sciences. Member ARS.

² Numbers in parentheses indicate References at end of paper.

³ In case of the turbulent boundary layer, ρ , u and T_0 refer to time-mean values.

r are interrelated by a constraint defining the generating curve of the surface.

The required integrals in Eqs. 2 and 3 are related to the stagnation temperature thickness (4)

$$\Gamma = \frac{1}{\rho_\infty U_\infty} \int_0^\infty \rho u \left(1 - \frac{T_0}{T_{0,\infty}}\right) dy \quad [4]$$

which is a measure of the amount of free stream flow required to account for the deficiency (or excess) of stagnation temperature flux in the free stream.

Introducing the thermal Reynolds number based on Γ

$$Re_\Gamma = \rho_\infty U_\infty \Gamma / \mu_\infty \quad [5]$$

we express Eqs. 2 and 3 in the form

$$q_{w,2} = \mu_\infty c_{p,\infty} T_{0,\infty} \frac{d}{ds} Re_\Gamma \quad [6]$$

$$q_{w,3} = \frac{\mu_\infty c_{p,\infty} T_{0,\infty}}{r} \frac{d}{ds} r Re_\Gamma \quad [7]$$

Finally, by introducing the definitions of the adiabatic heat transfer coefficient

$$h = q_w / (T_{ad} - T_w) \quad [8]$$

and the Stanton number

$$St = h / \rho_\infty c_{p,\infty} U_\infty \quad [9]$$

we convert the energy-integral equations into differential relations between Re_Γ and St in two-dimensional and axisymmetric flow

$$\frac{d}{ds} Re_\Gamma = \frac{\rho_\infty U_\infty}{\mu_\infty} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) St \quad [10a]$$

$$\frac{1}{r} \frac{d}{ds} r Re_\Gamma = \frac{\rho_\infty U_\infty}{\mu_\infty} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) St \quad [10b]$$

The adiabatic temperature T_{ad} is related to the stagnation and free stream temperatures by means of

$$\mathcal{R} = (T_{ad} - T_\infty) / (T_{0,\infty} - T_\infty) \quad [11a]$$

where \mathcal{R} is a semi-empirically determined recovery factor (5)

$$\mathcal{R} = \begin{cases} Pr^{1/2} & \text{laminar boundary layer} \\ Pr^{1/3} & \text{turbulent boundary layer} \end{cases} \quad [12]$$

In view of the fundamental aerothermodynamic relation (see 5)

$$T_{0,\infty} = T_\infty \left(1 + \frac{\gamma - 1}{2} M_\infty^2 \right) \quad [13]$$

the adiabatic wall temperature is, from Eq. 11a

$$T_{ad} = \frac{T_{0,\infty} \{ 1 + [(\gamma - 1)/2] M_\infty^2 \mathcal{R} \}}{1 + [(\gamma - 1)/2] M_\infty^2} \quad [11b]$$

where M_∞ is the free stream Mach number.

Solution of the Energy Integral Equations

It is readily seen that if the dependence of St on Re_Γ in Eqs. 10a and 10b were known, it would be a simple matter to solve these differential equations for Re_Γ and hence for h vs. s , in the presence of arbitrary free stream conditions, ρ_∞ , U_∞ , μ_∞ , vs. s .

The relation between Re_Γ and St can be readily determined for the particular case of uniform, longitudinal flow along a flat plate. In accordance with the method of (6), the particular functional relation thus obtained is assumed to hold more generally in the presence of variable ρ_∞ , U_∞ , μ_∞ , thus reducing the energy-integral equations (Eqs. 10a and 10b) to a first-order differential equation for Re_Γ or St . The validity

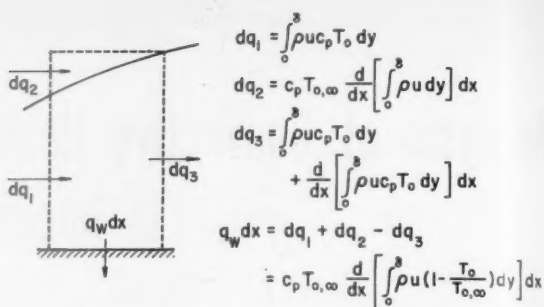


Fig. 1 Sketch of control volume for deriving energy-integral equation for boundary layer element of width dx , height δ see (5)

of such a procedure is supported by comparison of predicted h with test data for typical turbine blade profiles and flat plate data under both favorable and adverse pressure gradients (6).

In order to determine the relation between Re_Γ and St we start with the Blasius formula for the flat plate-skin friction coefficient for constant fluid properties⁴

$$C_{f,\infty}/2 = B Re_{x,\infty}^{-b} \quad [14]$$

and employ a modified Reynolds analogy

$$St_\infty = (C_{f,\infty}/2) Pr_\infty^{-2/3} = B Pr_\infty^{-2/3} Re_{x,\infty}^{-b} \quad [15]$$

where the semi-empirical Blasius parameters b , B have the values listed in Table 1.

Under the reference plate conditions with ρ_∞ , U_∞ , μ_∞ constant, the plate Reynolds number $Re_{x,\infty} = \rho_\infty U_\infty x / \mu_\infty$ is proportional to x with origin at the leading edge. Substituting for St_∞ from Eq. 15 into Eq. 10a and solving for the relation between Re_Γ and $Re_{x,\infty}$ one obtains

$$Re_\Gamma = \frac{B Pr_\infty^{-2/3}}{1 - b} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) Re_{x,\infty}^{1-b} \quad [16]$$

where the initial condition

$$Re_\Gamma = 0 \quad \text{at} \quad x = 0 \quad [17]$$

has been employed. From Eqs. 15 and 16, the reference relation between Re_Γ and St_∞ is, therefore

$$Re_\Gamma = \frac{(B Pr_\infty^{-2/3})^{1/b}}{1 - b} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) (St_\infty)^{-(1-b)/b} \quad [18]$$

Under the assumption that Eq. 18 is generally valid for variable free stream conditions ρ_∞ , U_∞ , μ_∞ , one obtains from Eqs. 10a and 10b the following first-order differential equations for the curved plate and the axisymmetric surface, respectively⁵

$$\frac{1}{St_\infty} \frac{d}{ds} (St_\infty)^{-(1-b)/b} = \frac{1 - b}{(B Pr_\infty^{-2/3})^{1/b}} \cdot \frac{\rho_\infty U_\infty}{\mu_\infty} \quad [19]$$

$$\frac{1}{r St_\infty} \frac{d}{ds} [r (St_\infty)^{-(1-b)/b}] = \frac{1 - b}{(B Pr_\infty^{-2/3})^{1/b}} \cdot \frac{\rho_\infty U_\infty}{\mu_\infty} \quad [20]$$

The solutions for St_∞ in terms of a generalized Reynolds number

$$Re_{s,\infty} = \int_0^s \frac{\rho_\infty U_\infty}{\mu_\infty} ds \quad [21]$$

⁴ The designation $c_{p,\infty}$, μ_∞ , Pr_∞ for the constant fluid properties is retained in Eqs. 11 to 23.

⁵ In obtaining Eqs. 19 and 20, the variation of $(T_{ad} - T_w)/T_{0,\infty}$ with s has been neglected, permitting cancellation of this factor, which is nearly constant in rocket application.

are obtained by elementary integration with the initial condition that St_∞ is infinite at the leading edge $s = 0$ (i.e. $Re_\Gamma = 0$ at $s = 0$). From the results thus obtained, the heat transfer coefficients h , related to St by the definition Eq. 9, are given by the following formulas

$$h_2(s) = BPr_\infty^{-2/3} \rho_\infty c_{p,\infty} U_\infty / Re_{s,\infty}^b \quad [22]$$

$$h_3(s) = h_2(s) \left[\frac{r^{1/(1-b)} Re_{s,\infty}}{\int_0^s r^{1/(1-b)} \rho_\infty U_\infty \mu_\infty^{-1} ds} \right]^b \quad [23]$$

Influence of Variable Fluid Properties

If fluid property variations due to temperature gradients in the boundary layer and compressibility effects associated with high velocity flows are taken into account, modified forms of Eqs. 22 and 23 are conveniently obtained by application of Eckert's reference temperature method (7). For the purpose of estimating rocket nozzle heat transfer coefficients, it is sufficient to consider density variation based on the ideal gas law

$$\rho \propto T^{-1} \quad [24]$$

and temperature-dependent viscosity expressed by

$$\mu \propto T^\omega \quad [25]$$

where ω is an empirically determined parameter (0.76 for air). In regard to the Prandtl number and specific heat, these properties show a relatively weak temperature dependence, and, for the sake of simplicity, they will be assumed constant at a mean value between $T_{0,\infty}$ and T_w designated by \bar{Pr} and \bar{c}_p , respectively. According to the method of (7), the modified Stanton number (and friction coefficient) for the flat plate with uniform free stream conditions can be evaluated from the following formulas

$$\frac{St}{St_\infty} = \frac{C_f}{C_{f,\infty}} = \left(\frac{\mu_\infty}{\mu^*} \right)^{-b} \left(\frac{\rho^*}{\rho_\infty} \right)^{1-b} \quad [26]$$

where the fluid properties μ^* and ρ^* are introduced at an empirically determined reference temperature (7)

$$T^* = 0.50(T_w + T_\infty) + 0.22(T_{ad} - T_\infty) \quad [27]$$

Thus with the property dependences in Eqs. 24, 25, the modified Stanton number for the flat plate becomes

$$St = \beta St_\infty = \beta BPr_\infty^{-2/3} Re_{s,\infty}^{-b} \quad [28]$$

where β is a function of the reference temperature

$$\beta = (T_\infty/T^*)^{1-b(1+\omega)} \quad [29]$$

Repeating the procedure of the previous section we now determine the flat plate reference relation between Re_Γ and St under variable fluid properties in the boundary layer represented by the factor β , which is constant for uniform longitudinal flow along the plate. Substitution of St from Eq. 28 into Eq. 10 now yields

$$Re_\Gamma = \frac{(\beta BPr_\infty^{-2/3})^{1/b}}{1-b} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) Re_{s,\infty}^{1-b} \quad [30]$$

and, upon elimination of $Re_{s,\infty}$ in favor of St

$$Re_\Gamma = \frac{(\beta BPr_\infty^{-2/3})^{1/b}}{1-b} \left(\frac{T_{ad} - T_w}{T_{0,\infty}} \right) (St)^{-(1-b)/b} \quad [31]$$

Assuming again that the reference flat plate relation, Eq. 31, holds generally under variable ρ_∞ , U_∞ , μ_∞ , and β one obtains from Eqs. 10a and 10b the differential equations for curved plate and axisymmetric surfaces, respectively

$$\frac{1}{St} \frac{d}{ds} \beta^{1/b} (St)^{-(1-b)/b} = \frac{1-b}{(BPr_\infty^{-2/3})^{1/b}} \cdot \frac{\rho_\infty U_\infty}{\mu_\infty} \quad [32]$$

Table 1 Semi-empirical Blasius parameters

Flow in boundary layer	b	B
laminar	$1/2$	0.332
turbulent	$1/5$	0.0296

$$\frac{1}{rSt} \frac{d}{ds} r \beta^{1/b} (St)^{-(1-b)/b} = \frac{1-b}{(BPr_\infty^{-2/3})^{1/b}} \cdot \frac{\rho_\infty U_\infty}{\mu_\infty} \quad [33]$$

where, as previously, the variation of the factor $T_{ad} - T_w / T_{0,\infty}$ with s has been neglected. It is convenient to consider

$$G = \beta^{1/(1-b)} St \quad [34]$$

as the dependent variable in Eqs. 32 and 33, which can then be readily solved for G upon separation of variables and integration. This procedure leads to the following results for the convective heat transfer coefficients for the curved plate and the corresponding axisymmetric surface

$$h_2(s) = \rho_\infty \bar{c}_p U_\infty St = \frac{\beta^{1/(1-b)} BPr_\infty^{-2/3} \rho_\infty \bar{c}_p U_\infty}{\left[\int_0^s \beta^{1/(1-b)} \rho_\infty U_\infty \mu_\infty^{-1} ds \right]^b} \quad [35]$$

$$h_3(s) = h_2(s) \left[\frac{r^{1/(1-b)} \int_0^s \beta^{1/(1-b)} \rho_\infty U_\infty \mu_\infty^{-1} ds}{\int_0^s (\beta r)^{1/(1-b)} \rho_\infty U_\infty \mu_\infty^{-1} ds} \right]^b \quad [36]$$

The generalization of these results to include in a better approximation the variations of Pr^* , c_p^* and $(T_{ad} - T_w) / T_{0,\infty}$ with s is straightforward but tedious, and does not appear warranted in application to rocket nozzle flow.

Summary of Procedure

For the estimation of convective heat flux in advanced rocket nozzles, the preceding developments require the following input data:

1 Wall configuration specified by $r = r(s)$, where s is the longitudinal (or meridional) coordinate with origin $s = 0$ at the injector, and r is the radius of rotation at the coordinate s ; wall temperature specified by constant T_w .

2 Combustion product data $T_{0,\infty}$, \bar{Pr} , \bar{c}_p , and γ ; and ω corresponding to the temperature dependence of viscosity in Eq. 25.

3 Free stream data ρ_∞ , U_∞ , M_∞ , vs. s as determined from the inviscid flow analysis of the exhaust stream.

The calculation procedure involves the following steps utilizing the Blasius parameters $b = 1/5$, $B = 0.0296$ and recovery factor $\mathcal{R} = \bar{Pr}^{1/3}$ corresponding to turbulent boundary layers:

1 Calculate the local adiabatic wall temperature T_{ad} vs. s from (See Eqs. 11a and 12)

$$T_{ad} = T_{0,\infty} \left(\frac{1 + [(\gamma - 1)/2] M_\infty^2 \bar{Pr}^{1/3}}{1 + [(\gamma - 1)/2] M_\infty^2} \right)$$

2 Calculate the local viscosity μ_∞ in accordance with the temperature dependence (see Eq. 25) $\mu_\infty \propto T_{0,\infty}^\omega$.

3 Calculate the reference temperature ratio T^*/T_∞ vs. s from Eq. 27, with the aid of Eq. 13.

4 Determine the function β vs. s based on (see Eq. 29)

$$\beta = \left(\frac{T_\infty}{T^*} \right)^{0.8-0.2\omega}$$

5 Calculate the curved plate heat transfer coefficient from (see Eq. 35)

$$h_2(s) = \frac{0.296 \beta^{1/4} \bar{Pr}^{-2/3} \rho_\infty \bar{c}_p U_\infty}{\left(\int_0^s \beta^{1/4} \rho_\infty U_\infty \mu_\infty^{-1} ds \right)^{1/5}}$$

6 Calculate the heat transfer coefficient for the corresponding axisymmetric nozzle from (see Eq. 36)

$$h_3(s) = h_2(s) \left[\frac{r^{5/4} \int_0^s \beta^{3/4} \rho_\infty U_\infty \mu_\infty^{-1} ds}{\int_0^s (r\beta)^{5/4} \rho_\infty U_\infty \mu_\infty^{-1} ds} \right]^{1/5}$$

7 Calculate the local heat flux from Eq. 8.

Table 2 Combustion product data for N_2O_4 /UDMH- N_2H_4 at $p_c = 300$ psia

$T_{0_{\infty}} = 5776$ R	$c_p = 0.45$ Btu/lb-deg R
$\mu_\infty(T_{0_{\infty}}) = 5.25 \times 10^{-6}$ lb/in.-sec	
$\bar{P}_r = 0.80$	$\gamma = 1.25$ $\omega = 1$

Table 3 Contour and free-stream data
a Outer contour

s , in.	r , in.	M_∞	ρ_∞ , lb/ft ³	U_∞ , ft/sec
0	10.28	0.250	0.1052	922
3.0	10.28	0.300	0.1038	1188
6.0	10.28	0.310	0.1034	1229
9.0	10.28	0.324	0.1030	1289
12.1	10.15	0.344	0.1022	1361
15.0	9.89	0.376	0.0996	1486
18.6	9.61	0.420	0.0993	1659
19.3	9.55	0.450	0.0981	1772
19.6	9.53	0.510	0.0956	1999
20.2	9.67	0.550	0.0937	2149
21.0	10.03	0.640	0.0888	2489
21.4	10.35	0.700	0.0849	2753
21.9	10.65	1.000	0.0676	3771
22.1	10.82	1.177	0.0574	4344
24.4	12.89	1.449	0.0428	5159
25.3	13.59	1.620	0.0350	5623
26.3	14.44	1.808	0.0276	6090
27.7	15.49	2.014	0.0210	6562
29.9	16.85	2.245	0.0154	7033
32.0	18.11	2.439	0.0118	7387
34.0	19.13	2.579	0.0097	7621
36.4	20.26	2.729	0.0078	7855
39.6	21.55	2.890	0.0062	8085
43.8	23.01	3.060	0.0049	8311
49.5	24.59	3.247	0.0039	8530
57.4	26.31	3.444	0.0029	8741
65.4	27.52	3.594	0.0023	8892

b Inner contour (plug)

s	r	M_∞	ρ_∞	U_∞
0	5.87	0.250	0.1052	992
3.0	5.87	0.300	0.1038	1188
6.0	5.87	0.310	0.1034	1229
9.0	5.87	0.324	0.1030	1282
12.0	5.87	0.344	0.1022	1361
15.0	5.87	0.376	0.0996	1436
18.5	5.87	0.420	0.0993	1659
19.88	6.11	0.450	0.0981	1772
21.33	6.70	0.510	0.0956	1999
22.89	7.72	0.550	0.0937	2149
23.33	8.93	0.640	0.0888	2489
25.05	9.60	0.700	0.0848	2753
25.68	10.17	1.000	0.0676	3771
25.96	10.43	1.177	0.0574	4344
25.99	10.67	1.651	0.0336	5706
26.12	10.78	2.096	0.0121	6743
26.25	10.87	2.589	0.0096	7636
26.38	10.90	3.154	0.0043	8421
26.52	10.93	3.845	0.0016	9111

Illustrative Application to an E-D Nozzle

The calculation procedure summarized above is now applied to a representative Rocketdyne E-D nozzle. The wall contours considered are sketched in Fig. 2. The heat transfer coefficient is calculated for a storable propellant system N_2O_4 /UDMH- N_2H_4 with combustion products at a chamber pressure of $p_c = 300$ psia. The required combustion product data are listed in Table 2. A wall temperature of $T_w = 2000$ R is assumed in the calculations. The contour specification and required free stream data obtained from the inviscid flow analysis is listed in Table 3.

The calculated results h_3 for the inner and outer contours are shown in Figs. 3b and 3a respectively. These curves show a characteristic rise toward the injector end, typical of boundary layer results at a leading edge. Close to the injector the calculation has, of course, no validity, but as noted in previous boundary layer investigations of rocket heat transfer (8,9), the starting conditions have negligible effect on the predicted heat transfer near the throat and in the supersonic region of the nozzle.

In regard to the plug contour, downstream of the "throat" where separation and recirculation occurs in practice, the predicted results (shown as a dashed curve in Fig. 3b) are not applicable.

Comparisons of Predictions for deLaval Nozzles

For the conventional deLaval nozzles the heat transfer formulas based on the boundary layer analysis permits a con-

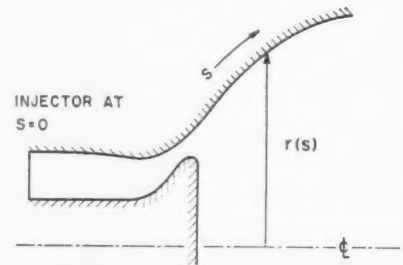


Fig. 2 Sketch of E-D thrust chamber configuration

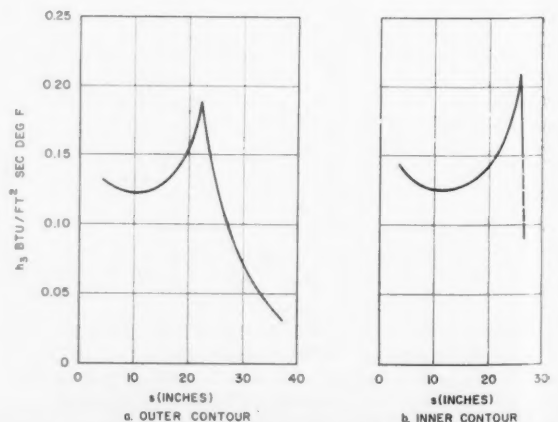


Fig. 3 Calculated heat transfer coefficient in E-D nozzle (N_2O_4 /UDMH- N_2H_4 Combustion Products at $p_c = 300$ psia)

venient parametric comparison with the conventional heat transfer formulas based on fully developed turbulent flow in tubes. We consider in particular the tube flow heat transfer correlation (3)

$$Nu_D = 0.023 Re_D^{0.8} Pr^{0.4} \quad [37]$$

where, as indicated by the subscript D , the Nusselt and Reynolds numbers are based on the local tube diameter D . Fluid properties required in Eq. 37 are generally introduced at the arithmetic mean between bulk and wall temperatures.

We shall now compare the prediction based on Eq. 37 with those obtainable from the boundary layer (Eqs. 22 and 23) at the throat $D = D_t$, where heat transfer conditions are most severe. In order to obtain meaningful comparisons we consider identical mass flows through a throat of fixed diameter D_t with chamber configuration parameters interrelated by constant values of L^* ; i.e., constant ratio of volume to throat area. For simplicity we consider the chamber as a combination of a cylinder (diameter D_c , length L_c) with a convergent cone (half-angle ϕ , contraction ratio $\epsilon_c = D_c^2/D_t^2$). Under these conditions the local free stream data are related to Re_D in Eq. 37 by

$$\rho_\infty U_\infty = \bar{\mu} Re_D / D \quad [38a]$$

and because of the constancy of $\rho_\infty U_\infty D^2$ (continuity of mass flow)

$$\rho_\infty U_\infty = \bar{\mu} Re_{D_t} D_t / D^2 \quad [38b]$$

Substituting from Eq. 38b into Eq. 22 with specific heat and viscosity at the mean values \bar{c}_p and $\bar{\mu}$, respectively, one obtains

$$h_{2t} = 0.0296 \bar{P} r^{-2/3} \bar{\mu} \bar{c}_p Re_{D_t}^{1/2} D_t^{-1} I^{-1/2} \quad [39]$$

where I is the chamber configuration integral from the injector $s = 0$ to the throat $s = s_t$

$$I = D_t \int_0^{s_t} D^{-2} ds \quad [40]$$

For the cylinder-cone configuration with the parameters ϵ_c, L_c, ϕ , the integral in Eq. 40 is

$$I = \frac{1}{\epsilon_c} \cdot \frac{L_c}{D_t} + \frac{1}{2 \sin \phi} \left(1 - \frac{1}{\epsilon_c^{1/2}} \right) \quad [41]$$

By further restricting the configuration to constant value of L^* , the ratio of volume to throat area, the parameter L_c can be eliminated from Eq. 41 by means of

$$\frac{L_c}{D_t} = \frac{1}{\epsilon_c} \left[\frac{L^*}{D_t} - \frac{\cot \phi}{6} (\epsilon_c^{3/2} - 1) \right] \quad [42]$$

Thus Eqs. 41 and 42 define I as a function of L^*, ϵ_c, D_t and ϕ .

The heat transfer coefficient h_3 at the throat of the deLaval nozzle is now calculated from Eq. 23 which leads to

$$h_{3t} = h_{2t} J^{1/2} \quad [43]$$

where J is the configuration factor

$$J = \frac{D_t^{5/4} \int_0^{s_t} D^{-2} ds}{\int_0^{s_t} D^{-3/4} ds} \quad [44]$$

For the cylinder-cone combination the indicated integrations yield

$$J = \frac{\epsilon_c^{-1/2} \frac{L_c}{D_t} + \frac{1}{2 \sin \phi} (\epsilon_c^{1/2} - 1)}{\epsilon_c^{1/8} \frac{L_c}{D_t} + \frac{2}{\sin \phi} (\epsilon_c^{5/8} - \epsilon_c^{1/2})} \quad [45]$$

which together with Eq. 42 define J in terms of the parameters $L^*, \epsilon_c, D_t, \phi$.

From the pipe flow correlation in Eq. 37 the heat transfer

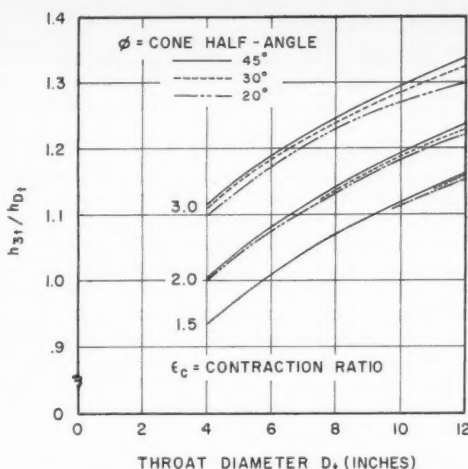


Fig. 4 Comparison of heat transfer coefficients calculated by two methods for chambers of $L^* = 50$ in.: h_{3t} (Boundary Layer) and h_{Dt} (Pipe Flow Correlation)

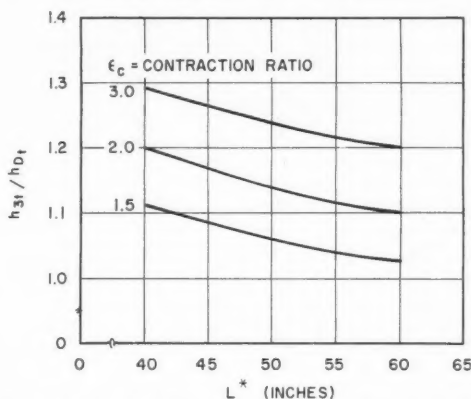


Fig. 5 Variation of h_{3t}/h_{Dt} with L^* for $D_t = 8$ in. $\phi = 30$ deg.

coefficient at the throat is

$$h_{Dt} = 0.023 \bar{k} D_t^{-1} Re_{D_t}^{0.8} Pr^{-0.4} \quad [46]$$

where \bar{k} is the mean thermal conductivity related to $\bar{P}r$ by

$$\bar{k} = \bar{\mu} \bar{c}_p / \bar{P}r \quad [47]$$

Therefore the ratio of h_{3t} predicted by the boundary layer formulas to that obtained from the pipe flow correlation is

$$\frac{h_{3t}}{h_{Dt}} = 1.28 \bar{P}r^{-0.07} \left(\frac{J}{I} \right)^{1/5} \quad [48]$$

For a variety of rocket combustion products the factor $\bar{P}r^{-0.07}$ is nearly constant and equal to about 1.02. Therefore, the ratio h_{3t}/h_{Dt} depends primarily on the chamber configuration parameters L^*, ϵ_c, D_t and ϕ . The parametric dependence⁶ is shown in Fig. 4 for representative value of $L^* = 50$ in. Curves of the ratio h_{3t}/h_{Dt} for a range of L^* 's are shown in Fig. 5, calculated with $D_t = 8$ in., $\phi = 30$ deg.

⁶ Recent nozzle heat transfer data obtained at the Jet Propulsion Laboratories of CIT are reported in (12). The data summarized in Fig. 9 of (12) are consistent with the trends described by Eq. 48 and represented in Figs. 4 and 5, which predict increasing throat heat flux with decreasing L^* or increasing ϵ_c , other parameters being held constant. In particular, the reference data reveal an increase of the order of 10% in the peak heat flux when L^* is decreased from 62.8 to 38 in. by decreasing the length without changing the throat and injector configurations.

Concluding Remarks

In conclusion we note some advantages possessed by the simplified boundary layer solutions both in regard to physical insight and ease of application to nozzle heat transfer problems.

The fundamental solution, Eq. 22, for the two-dimensional flow field exhibits the role of two factors in their influence on the heat transfer coefficient: 1 the free stream convection factor $\rho_\infty U_\infty$, which may undergo rapid spatial variations, and 2 a monotonically increasing factor Re_s^b ($b > 0$). Thus the heat transfer coefficient is determined primarily by the local convection term $\rho_\infty U_\infty$ with secondary effects arising from upstream conditions implicit in Re_s^b . The conversion from two-dimensional to axisymmetric heat transfer coefficient given by Eq. 23 is essentially a geometric transformation. In particular, for the case of uniform free stream condition in axisymmetric external flow along a cone with r proportional to s , Eq. 23 reduces to

$$\frac{h_3}{h_2} = \left(\frac{2-b}{1-b} \right)^b \quad [49]$$

For the laminar boundary layer with $b = 1/2$ the above formula yields $(h_3/h_2)_{\text{lam}} = (3)^{1/2}$, in agreement with the results of the classical Mangler transformation (5). For the turbulent boundary layer with $b = 1/5$ and uniform free stream conditions as per supersonic external flow along the cone, Eq. 49 yields $(h_3/h_2)_{\text{turb}} = (2.25)^{1/5}$ in agreement with the result calculated in (10). For internal flow along cones, the expressions for the ratio of heat transfer (and skin friction) coefficients deduced from the transformation, Eq. 23, depend on the local contraction or expansion ratio, as well as the Blasius parameter b , in agreement with the qualitative trends discussed in (11).

In regard to the comparison of predictions based on boundary layer analysis and on the pipe flow correlation, the following trends are evident in Figs. 4 and 5:

1 h_{31}/h_{D1} increases with increasing throat diameter D_t , contraction ratio ϵ_c and to some extent with increasing convergence angle ϕ .

2 h_{31}/h_{D1} decreases with increasing L^* .

3 Good agreement between the two methods of calculation $h_{31}/h_{D1} \sim 1$ is obtained over an extended range of parameters representative of typical rocket motors ($L^* \sim 50$ in., $\epsilon_c \sim 2$, $\phi \sim 30$ deg, $D_t \sim 6$ in.).

The trends described are predicted by the upstream configuration factor $(J/I)^{1/5}$ in Eq. 48 resulting from the boundary layer method of analysis. From the viewpoint of the boundary layer theory, the applicability of pipe flow formulas to nozzle heat transfer calculation implied by (3) is a matter of coincidence realized in representative nozzle configurations for $L^* \sim 50$ in. It is anticipated that with improved combustion processes (better mixing, more highly reactive propellants) which permit significant reductions in L^* , the reduced boundary layer development will result in increased heat transfer at the throat. For such chambers the conventional pipe flow formulas are likely to underestimate the gas-side heat transfer coefficients. On the other hand, the boundary layer formula is likely to predict low nozzle heat transfer coefficients for relatively long chambers in which fully developed turbulent flow is achieved upstream of the throat.

Finally we note that the calculation of heat transfer coefficients from Eqs. 35 and 36 in accordance with the procedure described in the text requires straightforward quadratures readily accomplished by elementary graphical or numerical methods. In application to advanced rocket nozzles, the relative simplicity of this approximate procedure possesses significant advantages over more "rigorous" and elaborate boundary layer analyses. The potential accuracy of more elaborate analyses in estimating heat transfer cannot be realized in practice because of deviations from the idealized exhaust gas-flow in regard to temperature and fluid com-

position due to injection and combustion patterns in the rocket chamber.

Acknowledgments

This investigation was supported by Rocketdyne, a division of North American Aviation. The writer wishes to acknowledge, in particular, helpful discussions with members of the Basic Studies unit in the Advance Design Division of Rocketdyne.

Nomenclature

b	= semi-empirical Blasius parameter
c_p	= gas specific heat at constant pressure
h	= convective heat transfer coefficient
k	= gas thermal conductivity
p_c	= combustion chamber pressure
q	= heat flux
r	= radial coordinate
s	= meridional coordinate
u	= longitudinal velocity in boundary layer
x	= coordinate along flat plate
y	= coordinate normal to wall
B	= Blasius parameter
C_f	= skin friction coefficient
D	= diameter
I	= configuration integral
J	= configuration integral
L_c	= length of cylindrical portion of chamber
L^*	= ratio of chamber volume to throat area
M	= Mach number
Nu	= Nusselt number
Pr	= Prandtl number
α	= recovery factor
Re	= Reynolds number
St	= Stanton number
T	= temperature (absolute)
T^*	= reference temperature
T_{ad}	= adiabatic wall temperature
T_o	= stagnation temperature in boundary layer
T_w	= wall temperature
U_∞	= free-stream velocity
β	= reference temperature function
γ	= ratio of gas specific heats
δ	= boundary layer thickness
ϵ_c	= contraction ratio
μ	= dynamic viscosity
ρ	= density
ϕ	= cone half-angle
ω	= viscosity-temperature exponent
Γ	= stagnation temperature thickness

Subscripts

2	= two-dimensional (curved plate)
3	= three-dimensional (axisymmetric)
s	= curved wall
t	= throat
x	= flat plate
∞	= free stream
D	= diameter

References

- 1 Rao, G. V. R., "The E-D Nozzle," *ASTRONAUTICS*, vol. 5, no. 9, 1960, p. 28.
- 2 Berman, R. and Neuffer, B., "Plug-Nozzle Flexibility," *ASTRONAUTICS* vol. 5, no. 9, 1960, p. 30.
- 3 McAdams, W. H., *Heat Transmission*, McGraw-Hill Co., N. Y., 1954.
- 4 Summerfield, M., "Jet Propulsion Engines," vol. XII of *High Speed Aerodynamics and Jet Propulsion*, Princeton University Press, Princeton, N. J., 1958, p. 493.
- 5 Shapiro, A. H., *Compressible Fluid Flow*, vol. II, The Ronald Press Co., N. Y., 1954.
- 6 Ambrok, G. S., "Approximate Solution of Equations for Thermal Boundary Layer with Variations in Boundary Layer Structure," *Soviet Physics, Technical Physics* vol. 2, no. 9, 1957, p. 1979.
- 7 Eckert, E. R. G., "Engineering Relations for Heat Transfer and Friction," *Transactions of ASME*, vol. 78, 1956, p. 1273.
- 8 Bartz, D. R., "An Approximate Solution of Compressible Turbulent Boundary Layer Development and Convective Heat Transfer in Convergent-Divergent Nozzles," *Trans. ASME*, Nov. 1955.

9 Cohen, C. B. and Reshotko, E., "The Compressible Laminar Boundary Layer with Heat Transfer and Arbitrary Pressure Gradient," NACA Rep. 1294, 1956.
10 Gazley, Jr., C., "Theoretical Evaluation of the Turbulent Skin Friction and Heat Transfer on a Cone in Supersonic Flight," General Electric Co. Rep. R49A0524, 1949.

11 Schlichting, H., *Boundary Layer Theory*, McGraw-Hill Book Co., N. Y., 1955, Ch XXIIc.

12 Welsh, Jr., W. E. and Witte, A. B., "A Comparison of Analytical and Experimental Local Heat Fluxes in Liquid-Propellant Rocket Thrust Chambers," Jet Propulsion Laboratory TR 32-43, California Institute of Technology, Feb. 1 1961.

Ignition Characteristics of Metals and Alloys

L. E. DEAN¹ and
W. R. THOMPSON²

Aerojet-General Corp.
Sacramento, Calif.

The ignition characteristics of engine structural metals and alloys as influenced by the composition and pressure of the ambient atmosphere are of immediate interest to propulsion design engineers. Tubular test sections of the stainless steel, cobalt and nickel alloys, besides aluminum, copper and titanium were resistance heated in controlled atmospheres of oxygen, carbon dioxide and an equal mixture of these gases. Tube and gas temperatures obtained were correlated with color motion picture coverage of the manner in which the tube heated and failed. Stainless steels and cobalt alloys ignited within the melting point range of each material. Nickel alloys did not ignite until the melting point was reached. The rate of combustion increased with oxygen content. Stainless steels with a high nickel content appear most suited for applications at high temperatures in an oxidizing atmosphere.

THE ADVENT of nuclear energy and rocket propulsion has focused attention on the need for a more complete understanding of the mechanisms involved in the high temperature oxidation and ignition of metals in comparatively short time periods. The factors of metal composition and fabrication history, metal temperature and the composition, pressure and superficial velocity of the ambient atmosphere must be considered in any systematic evaluation of such ignition processes. In selecting materials to resist rapid oxidation, however, due emphasis must be given to other necessary metal prerequisites, including high thermal conductivity and tensile strength, low density, high melting point, ease of fabrication, and availability at reasonable cost.

Oxidation Mechanisms

Most metals and alloys, when exposed to an oxidizing atmosphere, undergo a chemical reaction with the oxidizing component of the gas phase, forming an oxide film or scale on

the surface of the metal. The physical and chemical reactions controlling oxidation (e.g., the rate at which the oxide forms and the nature of its bond to the metal surface) are complex functions of a number of variables. These include not only the chemical composition and physical state of the metal surface, and the geometry of the metal object, but also environmental factors (i.e., the chemical composition and fluid-dynamical characteristics of the gaseous environment and the pressure and temperature of the entire system).

Those metals that form a nonporous oxide film which adheres tightly to the base metal are immune to further oxidation. Ignition of the metal will not normally occur unless this protective film is broken. Nickel and aluminum have this desirable characteristic. However, if the oxide film is porous, or does not adhere to the base metal, oxidation will occur. This results in destruction of the metal through complete conversion to the oxide form. Ignition occurs more readily for metals forming such films. Mild steel, molybdenum and titanium have these characteristics. For most common metals, under ambient atmospheric conditions, such oxidation rates are comparatively slow.

Oxidation is an exothermic reaction. For slow rates of oxidation (e.g., the rusting of steel) the heat of reaction at the surface of the metal is dissipated principally by conduction to the bulk of the metal and by convection to the atmosphere. The metal temperature thus remains essentially at the ambient level. If, however the rate of oxidation is so rapid that

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹Head, Physical Research Section, Research and Material Dept., Liquid Rocket Plant. Member ARS.

²Senior Engineer, Advanced Research Div., Azusa, Calif. Member ARS.

³Numbers in parentheses indicate References at end of paper.

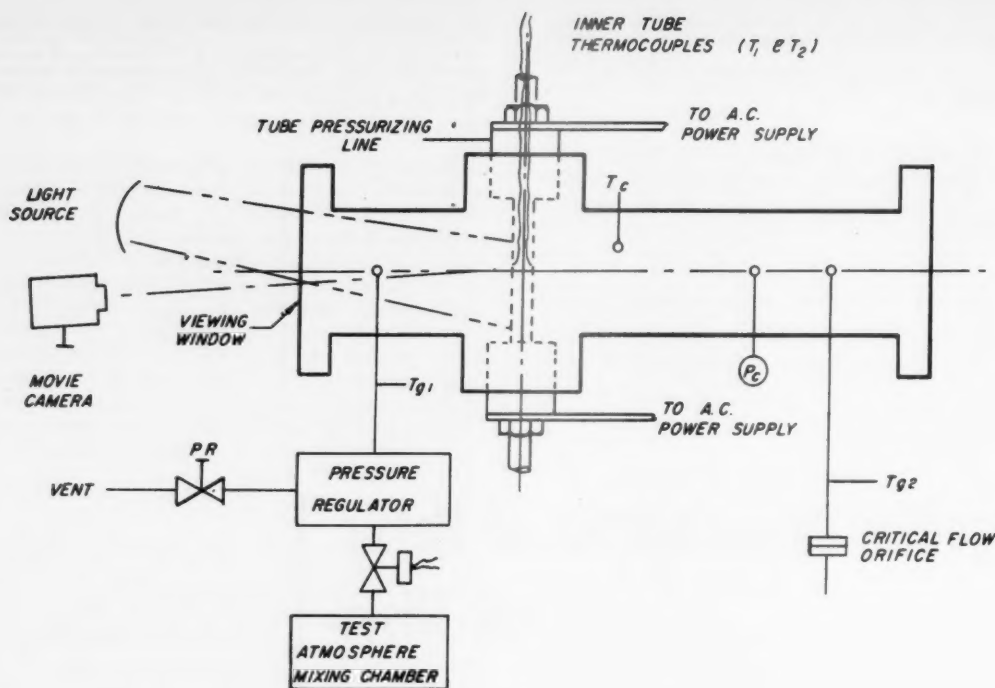


Fig. 1 Schematic of test apparatus

the generated heat cannot be removed from the surface at the same rate, the sensible heat of the surface increases with an accompanying temperature rise. This process continues until ignition of the metal occurs.

The problem of whether ignition, or kindling, of metals occurs below, at or above the melting point of the metal or alloy depends to an appreciable extent on the characteristics of the oxide film and on the overall rate of oxidation. For example, Fassell and co-workers (1) have shown that for magnesium alloys, the ignition temperature decreases as the oxidation rate is increased. The ignition temperatures of pure metals have been studied by Grosse and Conway (2) and by Mellor (3). Considerable data have been obtained by Reynolds (4) concerning the ignition of metals and various alloys in air and in oxygen and helium atmospheres under both stagnant and flowing conditions. In this investigation metal temperatures up to the point of ignition were measured by optical pyrometry. It was found that metals such as carbon steel, Type 410 stainless steel, molybdenum, tantalum, titanium and tungsten would ignite below their melting point. Metals such as Type 302 stainless steel and Inconel X ignited when the metal melted. No ignition occurred when nickel, Inconel and copper were tested in oxygen atmospheres. In general, the rate of burning was found to increase with pressure.

Hill and co-workers (5) reported that titanium, iron, carbon steel and 4130 steel spontaneously ignited in the solid phase (below the melting point) when heated in an atmosphere of oxygen. These metals melted rapidly while burning. Inconel, copper, 18-8 stainless steel, monel and aluminum could not be made to ignite spontaneously at temperatures up to melting with the equipment available. These tests were conducted under both static and flow conditions in air, oxygen and nitrogen atmospheres. Recent work at Aerojet-General (6) on the rates of self-sustaining burning of magnesium in

oxygen (with oxygen pressure and superficial velocity as variables) indicated that the burning rate followed either increased oxygen velocity or increased pressure. The rate controlling step appeared to be the maximum rate of vaporization of the metal at a temperature slightly above its boiling point. The same work contains a critical review and reinterpretation of the Eyring-Zwolinsky theory of metal ignition, based on the experimental data reported in (1). Microscopic intergranular oxidation at the surface of the metal, producing localized hot spots at the metal surface, may result in metal vaporization and ignition.

Description of the Test Apparatus and Testing Procedure

The test technique consisted basically of resistance-heating to destruction tubular specimens of various metals and alloys while surrounded by the test atmosphere at the required pressure. A schematic of the test apparatus is shown in Fig. 1. Tubular sections of the metal to be tested were clamped in the electrodes in a horizontal position in the testing device, parallel to the viewing port at one end of the cylinder. The chamber was pressurized to 50, 300 or 800 psia with either commercial-grade oxygen, carbon dioxide or a 50-50 mixture (by volume) of the two gases. Controlled cross flow of the chamber gas around the test piece to the outside air prevented stagnation of the atmosphere within the apparatus. The metal was then resistance heated by applying alternating current in step-wise increments at a rate so that metal failure occurred within one to two minutes after the initial application of power. Tube and gas temperatures, as well as power inputs, were continuously recorded. Color motion picture photography recorded the manner in which the tube heated and failed.

Test sections were 4 in. lengths of commercial tubing homo-






TEST ATMOSPHERE	
100% O ₂	
50% O ₂ - 50% CO ₂	
100% CO ₂	
100% NITROGEN	
100% HELIUM	

Fig. 2 Effect of atmosphere on kindling of 347 stainless steel (300 psia)

geneous to visual inspection. Tube diameters tested were either 0.500 or 0.375 in., whereas wall thicknesses varied from 0.010 to 0.035 in. It was necessary to machine some samples to obtain wall thicknesses having the desired electrical resistivity for cases where commercial tubing of the desired wall thickness was not available. Two no. 30 gauge chromel-alumel thermocouples were spotwelded or otherwise affixed to the inner tube walls at the midpoint of the tube (platinum-platinum 10% rhodium thermocouples were also used for a limited number of tests). Thermocouple lead wires were led through a 0.25-in. steel tube which was then filled with an epoxy resin to seal the outlet. A bleed orifice calculated to allow the desired gas cross-flow velocity around the test piece was installed downstream of the test chamber. The test atmosphere, prepared in the mixing chamber by pressurizing with the desired gases (using the pressure ratio as indicative of the volume fraction of each gas present), was then admitted to the test fixture at the desired pressure.

Power, controlled through a variable autotransformer, was applied to steel and alloy tubes in increments of approximately 0.25 v at 5-sec. intervals and increments of approximately 0.02 v at 5-sec intervals for the nonferrous test sections. Tests were continued until either tube failure occurred or the maximum voltage output of the transformer was reached. Two hundred individual tests were conducted in this manner.

Experimental Results

Two general series of tests were conducted. Since Type 347 stainless steel has wide application in fabricating propulsion system hardware, the ignition characteristics of this alloy were most intensively studied. More limited testing was performed on samples of the other metals and alloys.

347 Stainless Steel

A complete series of tests were conducted at 50, 300 and 800 psia in atmospheres of oxygen, carbon dioxide and an equal-volume mixture of these gases. Additional tests were made using nitrogen and helium atmospheres at the 300 psia pressure level. Tests were also conducted on the effects of superficial gas cross-flow velocity, rate of heating, and special metal surface treatment.

Effect of gas composition at 300 psia pressure. Failure of the test specimens occurred within the melting point range for all the gas compositions tested. The metal ignited in 100% oxygen and 50% O₂-50% CO₂ atmospheres with partial destruction of the test sections. Test results are shown in Fig. 2.

Effect of gas pressure. No ignition occurred in oxygen at 50 psia. At 300 and 800 psia ignition took place within the melting point range of the alloy, resulting in destruction of the test specimen. The rate of burning and proportion of test section destroyed increased with pressure. No ignition occurred in a carbon dioxide atmosphere. Post-test specimen appearance is shown in Fig. 3.

Effect of other factors. Wide variations in gas cross-flow velocity and in the rate of heating showed that ignition again took place within the melting point range of the metal. In tests where the tube was partially coated with a refractory ceramic, the metal ignited under the coating and burned as the coating flaked off the surface.

Carbon Steel and Other Stainless Steel Alloys

Effects of variations in gas composition and pressure upon the ignition of mild steel and Types 304, 310, 321, 410, AM350 and 17-7 PH stainless steel, comprises the bulk of the steel alloys used in rocket propulsion hardware.

Effect of an oxygen atmosphere. At a pressure of 50 psia, specimen ignition occurred within the melting point range of the metal with exception of Type 410 stainless steel which ignited at a lower temperature. At 800 psia, Type 304 stainless steels ignited at temperatures within the melting point range of the alloy. Ignition of the other metals tested oc-



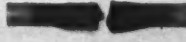






TEST ATMOSPHERE			
TEST PRESSURE	100% O ₂	50% O ₂ — 50% CO ₂	100% CO ₂
50 PSIA			
300 PSIA			
800 PSIA			

Fig. 3 Typical test results, 347 stainless steel

TEST ATMOSPHERE	100% O ₂	50% O ₂ — 50% CO ₂	100% CO ₂
CARBON STEEL			
17-7 PH			
310 S.S.			
410 S.S.			

Fig. 4 Effect of atmosphere on kindling of iron alloys (300 psia)

TEST ATMOSPHERE	100% O ₂	50% O ₂ — 50% CO ₂	100% CO ₂
NICKEL "A"			
INCONEL "X"			
MONEL			
HASTALLOY "X"			

Fig. 5 Effect of atmosphere on kindling of nickel alloys (300 psia)

occurred at temperatures of 250 to 500 F below their respective melting points. The rate of burning was found to increase with pressure. At 50 psia the carbon steel sample was completely destroyed, whereas about 60 to 80% of the stainless steel test sections were consumed. At the higher pressure levels, essentially all the tubing between the electrodes was destroyed.

Effect of gas composition at 300 psia pressure. Comparison of photographic and instrument data for tubes heated to destruction in 100% O₂, in 100% CO₂, and in an equal-volume mixture of these gases showed that, with the exception of Type 410 stainless steel, each test sample failed at a temperature within the melting point range of the alloy. In an oxygen atmosphere up to 70% of each test section was destroyed due to ignition and combustion of the metal. In the 50% O₂-50% CO₂ mixture, when ignition did occur with carbon steel and stainless types 347, 410 and AM 350 alloys, only partial destruction of the test specimen resulted. No ignition occurred in the carbon dioxide atmosphere; failure of the tube was due to melting. Test results are shown in Fig. 4.

Nickel Based Alloys

Test sections of six nickel based alloys were tested to destruction. These were Nickel A, Monel, Inconel X, Hastelloy C, Hastelloy R and Hastelloy X.

Effect of an oxygen atmosphere. At the 50 psia pressure level, test sections of these alloys all failed at temperatures within their melting point ranges. However, at 800 psia, both Inconel X and Monel failed at temperatures 250 to 500 F below their respective melting points. No ignition of the metal occurred with Nickel A at 50, 300 or 800 psia. Test sections of Monel, Inconel X, Hastelloy R and Hastelloy X ignited at a gas pressure of 800 psia. The ignition of Inconel X and Monel at the high gas pressure, and of Hastelloy R and Hastelloy X at all pressure levels resulted in partial destruction of the test section.

Effect of gas composition at 300 psia pressure. All test specimens failed at temperatures within their respective melting point ranges. Ignition of Hastelloy R in an oxygen atmosphere resulted in 60% destruction of the test section.

Although ignition occurred in oxygen for samples of Inconel X, Hastelloy C and Hastelloy X, only the portion of the tube immediately adjacent to the severance area was consumed. With the exception of Hastelloy R, no ignition occurred in the 50% O₂-50% CO₂ atmosphere. The appearance of these sections following the tests is shown in Fig. 5.

Cobalt Based Alloys

Two cobalt based alloys were studied in this program. These were Haynes 25 and Multimet.

Effect of an oxygen atmosphere. Ignition of test samples of these two alloys occurred within their respective melting point ranges. The rate of tube destruction increased with gas pressure; at the 50 psia pressure level, 5 to 30% of each test section was consumed after ignition, whereas at higher pressures 50 to 70% of each sample was destroyed.

Effect of gas composition at 300 psia pressure. Metal ignition occurred for both alloys in 100% oxygen with over 70% of each section destroyed. Test specimens of Multimet ignited in the 50-50 mixture of oxygen and carbon dioxide. The remains of the test samples are shown in Fig. 6.

Miscellaneous Metals

Of the nonferrous metals of importance to propulsion system design, aluminum, copper and titanium were selected for testing.

Effect of an oxygen atmosphere. Failure of the aluminum test section was due to melting of the metal with no subsequent ignition. At an oxygen pressure of 50 psia, no ignition occurred for the copper sample; however, at 300 psia ignition occurred and 60% of the test section was destroyed. Ignition of titanium occurred 250 to 1000 F below the melting point of the metal. Test sections were completely destroyed. Even the copper electrodes which held the test section in place

ignited and burned during the test conducted at 300 psia. Titanium was by far the most reactive of all the metals tested.

Effect of gas composition at 300 psia pressure. Failure of the aluminum test sections in oxygen, carbon dioxide and their equal-volume mixture was due solely to melting and subsequent severance of the tube. For copper, ignition and partial destruction was noted for tests conducted in oxygen with no ignition in other gas compositions. Titanium was observed to ignite in all the atmospheres tested. Even in the 100% CO₂ atmosphere the test specimen was completely destroyed at a temperature 250 F below the melting point of the metal. The remains of the samples after testing are shown in Fig. 7.

Discussion of Test Results

Carbon Steel and Stainless Steel Alloys

The ignition temperature of carbon steel in an oxygen atmosphere as reported in (3 and 5) varies from 1700 to 2300 F. In this study, the ignition temperature was found to vary between 2000 to 2700 F, depending upon the pressure. It is probable that the oxide film on the tube's surface, as noted in the motion picture coverage, influenced the ignition characteristics of the test piece.

Stainless steels having an appreciable nickel content may be raised to their melting point before ignition will occur. Alloys such as Type 410 and Type 430, which have no nickel as an alloying element, ignite at temperatures 250 to 300 F below their nominal melting points.

For all the ferrous alloys investigated, the rate of burning increased with increases in gas pressure. Since combustion depends upon a continuous renewal of reactants in the combustion zone, it is obvious that increasing the total oxygen pressure results in a proportionately greater concentration of oxygen molecules available to the flame front. After initial







TEST ATMOSPHERE	100% O ₂	50% O ₂ — 50% CO ₂	100% CO ₂
HAYNES 25			
MULTIMET			

Fig. 6 Effect of atmosphere on kindling of cobalt alloys (300 psia)

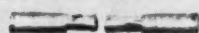
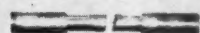







TEST ATMOSPHERE	100% O ₂	50% O ₂ — 50% CO ₂	100% CO ₂
ALUMINUM			
COPPER			
TITANIUM			

Fig. 7 Effect of atmosphere on kindling of metals (300 psia)

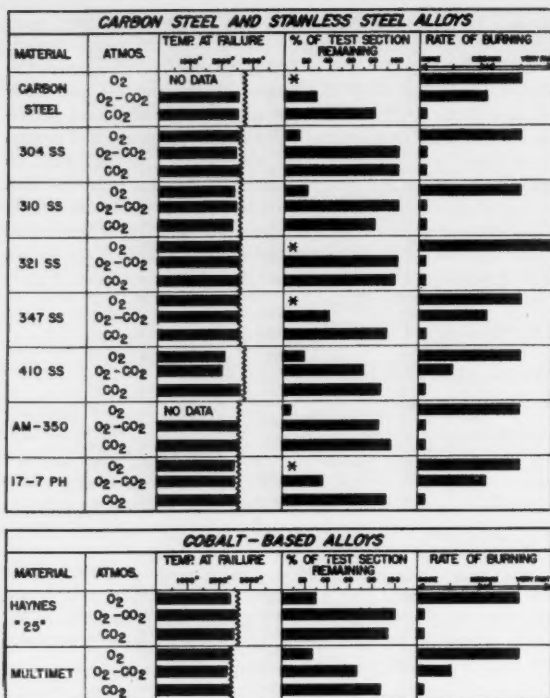


Fig. 8 Test summary in various atmospheres, pressure = 300 psia

ignition, the entire test section was destroyed at the higher pressure levels, whereas the test section was only partially destroyed for low pressure tests.

In atmospheres composed of 50% oxygen-50% carbon dioxide, ignition occurred within the melting point ranges of the particular metals tested (with the exception of Type 410 stainless steel). The rate of burning and the related length of sample consumed during the test decreased with decreasing oxygen concentration. Failure of the test sections in a carbon dioxide atmosphere occurred owing to melting and severance of the tube. No ignition was noted.

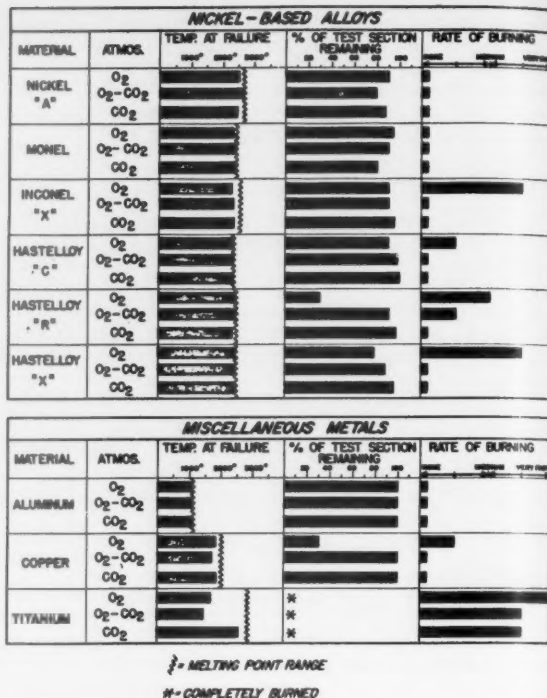
Nickel Based Alloys

Test results with Nickel A showed tube failure occurred by tube melting and severance. No ignition or combustion was observed in any test. Reynolds (4) reports similar results with Inconel in an oxygen atmosphere. Apparently a protective nonporous adherent oxide coating forms on the surface of the metal, inhibiting further oxidation.

Inconel X was found to ignite and burn at the 300 and 800 psi pressure levels. Only at the higher pressure did combustion completely destroy the tube. Reynolds also reports ignition and burning with Inconel X in an oxygen atmosphere. The ignition of Inconel X may possibly be attributed to the presence of 2-1/2% titanium in this alloy.

Of the three Hastelloy alloys evaluated, Hastelloy C appears most resistant to oxidation at high temperatures and pressures. Ignition and partial destruction of test sections of Hastelloy R and Hastelloy X occurred at both 300 and 800 psi gas pressure.

No ignition and subsequent combustion of nickel based alloys was noted in either carbon dioxide or 50% oxygen-50% carbon dioxide atmospheres with one exception. Partial destruction of a test section of Hastelloy R was observed for tests in the 50-50 mixture. Again, the ignition may result from the presence of titanium in the alloy.



Cobalt Based Alloys

In an oxygen atmosphere, ignition of the two cobalt based alloys occurred when the metal reached the melting point. The rate of destruction of the test section increased with increasing pressure, with both metals igniting almost explosively at the 800 psia level.

No ignition was observed for Haynes 25 in either carbon dioxide or the oxygen-carbon dioxide mixture at 300 psia, tube failure occurring due to melting. Multimet N-155 melted in the CO₂ atmosphere, whereas partial destruction occurred in the test with the O₂-CO₂ mixture.

Nonferrous Metals

No ignition or burning of the aluminum alloy occurred. Grosse and Conway (2) have shown that aluminum will not ignite until its temperature is raised above the melting point (1800 F). In the present study, evaluation of photo coverage showed that the tube melted and severed with no incandescence or hot spots. The high thermal conductivity allowed heat from the melting zone to be quickly dissipated, lowering the temperature below its kindling point.

Results of this study show that copper ignites at a temperature slightly below its melting point in a 100% oxygen atmosphere and at a pressure of 300 psia, resulting in the destruction of approximately 70% of the test specimen. Ignition of copper may be related to the condition of the oxide film and consequently dependent on the test conditions.

The results indicate that in an oxygen atmosphere, ignition of titanium will take place when the metal temperature exceeds 1500 F. In a carbon dioxide atmosphere, ignition was found to occur at a temperature 250 F below the melting point. It is reported (8 and 9) that titanium will react with carbon dioxide and form titanium dioxide and carbon monoxide. This reaction will occur at a temperature of 1650 F. Reynolds (4) also reports a wide range of temperatures wherein titanium

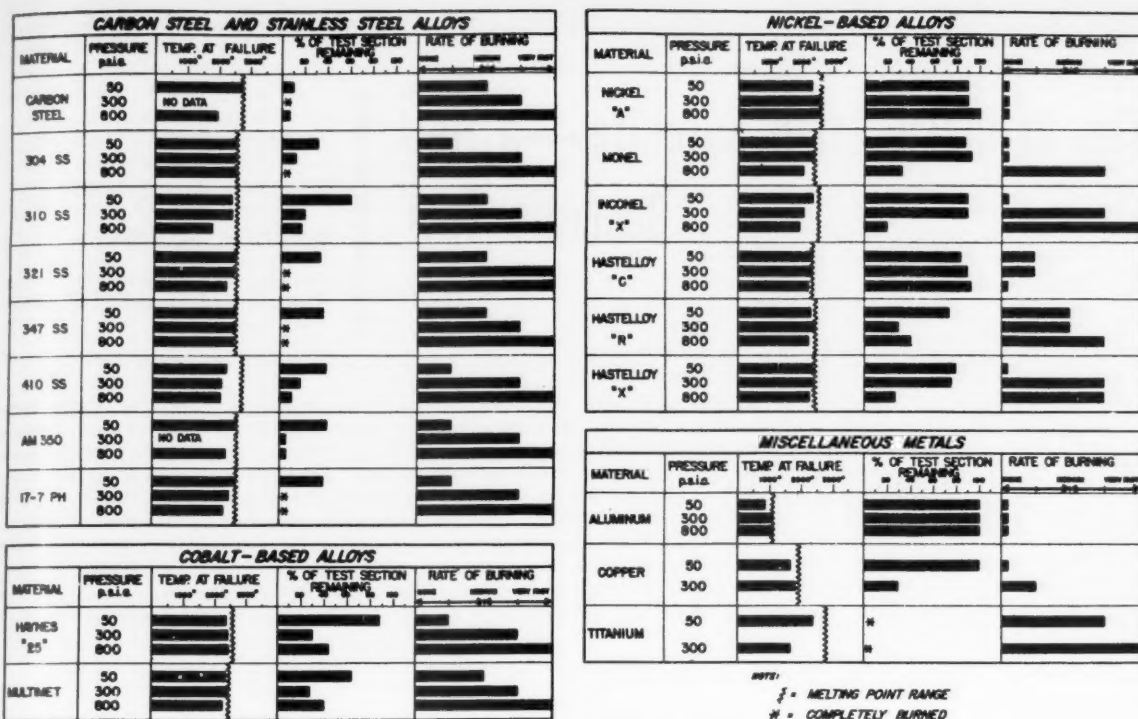


Fig. 9 Test summary in oxygen atmosphere

may ignite. These results indicate that ignition of titanium is also dependent upon the condition of the oxide film and consequently upon the particular testing technique employed. Of all the metals tested, titanium proved to be the most reactive metal at elevated temperatures regardless of the composition of the atmosphere.

Conclusions

The results of this investigation are graphically summarized in Fig. 8, in which the effect of gas composition at a constant pressure is shown, and in Fig. 9, in which the effect of pressure upon tubes in an oxygen atmosphere is noted. The following general conclusions are indicated.

The ignition of stainless steels (containing nickel) occurred within the melting point ranges of the alloys during tests in an oxygen atmosphere. Steel alloys, with no nickel content ignited at temperatures below their melting points. The rates of burning and destruction of the test sections increased with increasing oxygen concentrations.

Most of the nickel based alloys did not ignite until the melting point was reached. Nickel A did not ignite in oxygen, and, whereas ignition was noted for the other alloys at elevated pressures, the rate of burning of Nickel A was generally less severe than that observed for the ferrous based alloys.

Cobalt based alloys ignited at temperatures within their melting point range. The rates of burning were similar to those of the iron based alloys; however, the resulting damage to the test sections was less severe.

Ignition did not occur with the aluminum alloy studied. A copper test section ignited when tested in oxygen at the 300 psia pressure level. Titanium was the most reactive of all the metals tested, and was the only metal which ignited in an atmosphere of carbon dioxide.

From these test results it is concluded that stainless steels with a high nickel content are the most suitable material for the manufacture of such items as rocket engine combustion chambers. This metal is especially adaptable where high temperature environment plus resistance to oxidation reactions are major controlling factors.

References

- 1 Fassell, W. M., Gulbransen, L. B., Lewis, J. R. and Hamilton, J. H., "Ignition Temperatures of Magnesium and Magnesium Alloys," *J. Metals*, vol. 3, March 1940, pp. 522-528.
- 2 Grosse, A. V. and Conway, J. B., "Combustion of Metals in Oxygen," *Ind. and Eng. Chem.*, vol. 50, April 1958, pp. 663-672.
- 3 Mellor, J. W., *Treatise on Inorganic and Theoretical Chemistry*, Longmans, Green and Co., N. Y., 1946.
- 4 Reynolds, W. C., "Investigation of Ignition Temperatures of Solid Metals," NASA Tech. Note TN D-182, Oct. 1959.
- 5 Hill, P. R., Adamson, D., Fuland, D. H. and Bressette, W. E., "High-Temperature Oxidation and Ignition of Metals," NACA Rep. L55L23b, Research Mem. RM, March 1956.
- 6 "The Reaction of Metal in Oxidizing Gases at High Temperatures," Aerojet-General Rep. AGC-IDO-28000, Azusa, Calif., April 30, 1957.
- 7 Harrison, P. L., "The Combustion of Titanium and Zirconium," *Seventh Symposium (International) on Combustion*, Butterworths Scientific Publications, London, Eng.
- 8 Sidgwick, N. V., *The Chemical Elements and Their Compounds*, Oxford University Press, N. Y., 1950.
- 9 Moeller, T., *Inorganic Chemistry*, John Wiley and Sons, N. Y., 1952.

The Artificial Earth Satellite—a New Geodetic Tool¹

BRUCE C. MURRAY²

Division of Geological Sciences
California Institute of Technology
Pasadena, Calif.

An accurate satellite positioning system appropriate for geodetic applications is described. Observations could be made primarily by radar ranging, or possibly, by Doppler techniques. A flashing strobe light designed to be used in conjunction with existing Baker-Nunn and ballistic cameras might also be carried, particularly to provide an independent check of the electronic measurements. A 500 mile perigee height and low eccentricity orbit for the satellite is suggested. Higher satellite altitudes serve only to degrade observational accuracy. Irregular fluctuations in drag are believed to be sufficiently small at that height to permit accurate interpolation along the orbit from observations on one datum to those of another, unrelated datum. Gravitational perturbations become more important as the desired accuracy of the orbital interpolation increases; however, it is felt that such perturbations can be determined—and hence removed—by special analyses of observations of the geodetic satellite itself.

THE GEODETIC applications of artificial Earth satellites have been widely acknowledged since the onset of the International Geophysical Year when satellites were first seriously considered as geophysical tools. Yet, virtually no improved positional data of significance have been reported even at this date. Geodesists and other geophysicists did not fully appreciate the difficulties involved in using the artificial Earth satellite to improve the positions of the observing stations. Nor did they fully anticipate, on the other hand, the remarkable sensitivity of the artificial satellite as a tool for studying the gravity field of Earth. The reason for this disparity is that some gravity effects are accumulated over a number of revolutions in the form of anomalous orbital motions, whereas positional errors of the observing stations are difficult to distinguish from errors of the same order in the satellite observations themselves. Geodetic satellite applications thus have proceeded only slowly because the first few years of direct investigation of near space have brought to light a surprising complexity of nonconservative forces and irregularities in the potential fields in what have been generally thought of as empty, cold, and above all, simple space. However, enough has been learned in the last few years about this not-so-simple environment to once again warrant optimism concerning the improvement of station positions by a carefully conceived geodetic satellite system. It is the optimum design and application of such a system that is the subject of this paper.

General Survey of the Problem

From the foregoing, it is clear that improved station positions have not been a byproduct of general satellite tracking programs up to the present. Special efforts evidently are necessary in order to use an artificial Earth satellite as a precise geodetic tool. The problem, thus, is one of

optimization of observation parameters and of the observing stations. Stating the problem in this form seems, at first glance, to raise more questions than it answers. "Observation parameters" is a pleasant sounding cover-up for unruly questions concerning the accuracy and frequency of observations, concerning station locations, and concerning what ground-based and satellite-borne equipment should be used. This last question alone opens up many possible observing schemes because each of the three position coordinates and the three velocity coordinates of the satellite (with respect to the observer) can be measured both electronically and optically with varying degrees of success.

It is probable that only stellar photographic techniques for absolute angle measurement, radar ranging for distances, and radial velocity measurements by Doppler frequency shift at radio frequencies offer much hope of sufficient precision for use with a geodetic satellite in the immediate future. Yet, there is little general agreement as to (a) the optimum balance of optical and radar measurements, (b) the relative merits of C. W. and pulse radar systems of precise ranging, and (c) the usefulness of presently developed Doppler radial velocity systems for use with a geodetic satellite.

The situation regarding the optimum orbital parameters is also clouded. Some geodesists wish to use a satellite in an intervisible mode; i.e., in a simple extrapolation of the method of flare triangulation so that the satellite is simultaneously observed from a group of observation stations as illustrated in Fig. 1. Such simultaneous observations lead to a determination of relative orientation and position of the observing stations independent of any assumptions regarding either orbital theory or the nature of Earth's gravity field. Other geodesists feel that greater precision can be obtained by observing the satellite under the most favorable geometrical circumstances and relying on the orbit as an interpolation device from one set of observations to another, noninter-visible, set as illustrated in Fig. 1. Inasmuch as the two methods are incompatible in any one particular orbit, a choice must be made if that satellite is to be of optimum use in either method. Less controversy has arisen regarding the optimum inclination angle, but suggestions have ranged from 50–90 deg.

Presented at the ARS Semi-Annual Meeting, Los Angeles, Calif., May 9–12, 1960.

¹ This work was carried out while the author was a geophysicist in the Geophysics Research Directorate, Air Force Cambridge Research Laboratories, Bedford, Mass.

² Research Fellow in Space Science, Div. of Geological Sciences, California Institute of Technology, Pasadena, Calif.

The problem of optimization of observations and orbital parameters to improve station coordinates thus leads to a myriad of possible solutions without order or consistency. It is evident that some ground rules are needed if practical recommendations are going to be formulated on how to obtain geodetic data from satellite observations. Such ground rules are readily obtained from careful analysis of the geodetic objectives of a geodetic satellite. This is the procedure to be followed in the present paper, after a brief discussion of geodetic coordinate systems and their accuracies.

The Geodetic Problem and Associated Coordinate Systems

Geodesy traditionally has been concerned with the precise position of control points located on the surface of Earth. These are the building blocks from which the scientifically significant quantities, size, shape, and gravity field of Earth, are derived. Ideally, such control points would be referenced to a rotating system of Cartesian (X, Y, Z), spherical (ρ, ψ, λ), or ellipsoidal (R, ϕ, λ, e, a), coordinates centered at the exact center of mass of Earth with one axis coincident with the mean polar axis of Earth and with another lying in the zero longitude meridian plane.³ Such systems are equivalent forms of the geocentric coordinate system. However, it has not been possible in practice to reference surface positions to the center of mass of Earth. Consequently, two major departures from the geocentric coordinate system have been necessarily introduced in practice:

1 Horizontal distances and angles between triangulation stations are adjusted, ideally, with reference to a fixed initial point ϕ_{G0}, λ_{G0} and initial reference azimuth A_{G0} . (Let us assume, for discussion, that all three initial quantities have been obtained astronomically.) A reference ellipsoid of given semimajor axis and flattening is used as a computational model for reduction of the planar measurements to a known curvilinear system. The resulting geodetic coordinates ϕ_G, λ_G of the control points are thus referenced to a geodetic datum and serve to specify which particular normal of the reference ellipsoid each control point corresponds to. The

geodetic axes are held parallel to the geocentric set with high precision by frequent astronomical determination of the geocentric direction of the local vertical. But the initially assumed geodetic coordinates ϕ_{G0}, λ_{G0} of the datum differ from the true geocentric coordinates ϕ, λ of that same point by some small but uncertain amount, because the centers of the geodetic and geocentric ellipsoids are not coincident. Accordingly, two small displacements $R\Delta\phi_{G0}$ and $R\Delta\lambda_{G0}$, plus a small correction, ΔA_{G0} , to the reference azimuth are sufficient to bring the set of geodetic coordinates ϕ_G, λ_G of the control points into complete agreement with the set of geocentric values ϕ, λ of the same set of points, assuming the same eccentricity e is used for the geocentric ellipsoid. Only two of the three position coordinates of the points on a datum are defined by ϕ_G and λ_G , however, and for that reason a datum is not a complete coordinate system. A third coordinate, related to H , is required for each point before we may properly speak of a complete geodetic coordinate system.

2 The leveled height h is referenced to the geoid, the equipotential surface corresponding to mean sea level. But this vertical reference surface is an irregularly undulating surface. The slope of the geoid with respect to a datum is obtained by comparison of the measured geocentric direction of the local vertical at stations ϕ_G, λ_G with the computed direction of the normal to the reference ellipsoid at those stations. The resulting map of the height N_G of the geoid above the reference ellipsoid is obtained by integration of the slope values. However, inasmuch as this map is based on normals to the reference ellipsoid rather than to the geocentric one, there are systematic differences of the set N_G as compared to the set N . These differences are eliminated by the translation $R\Delta\phi_{G0}, R\Delta\lambda_{G0}$ which brings the geodetic and geocentric normals into coincidence. Also, ΔA_{G0} , the correction to the initial astronomic azimuth, is obtained directly once the true geocentric longitude has been obtained. Consequently, only a third simple translation ΔR of the entire geoid along the normal ϕ_{G0}, λ_{G0} is necessary to complete the transformation of geodetic coordinates into the geocentric system.

Thus, a set of geodetic positions ϕ_G, λ_G referenced to a geodetic datum, plus leveled heights h and computed geoid heights N_G , can be viewed as a particular ellipsoidal coordinate system which differs from a geocentric ellipsoidal system, of the same assumed a and e , by three small displacements.

³ For the excellent discussion of terrestrial coordinate systems and problems of three-dimensional geodesy, see Veis (1).

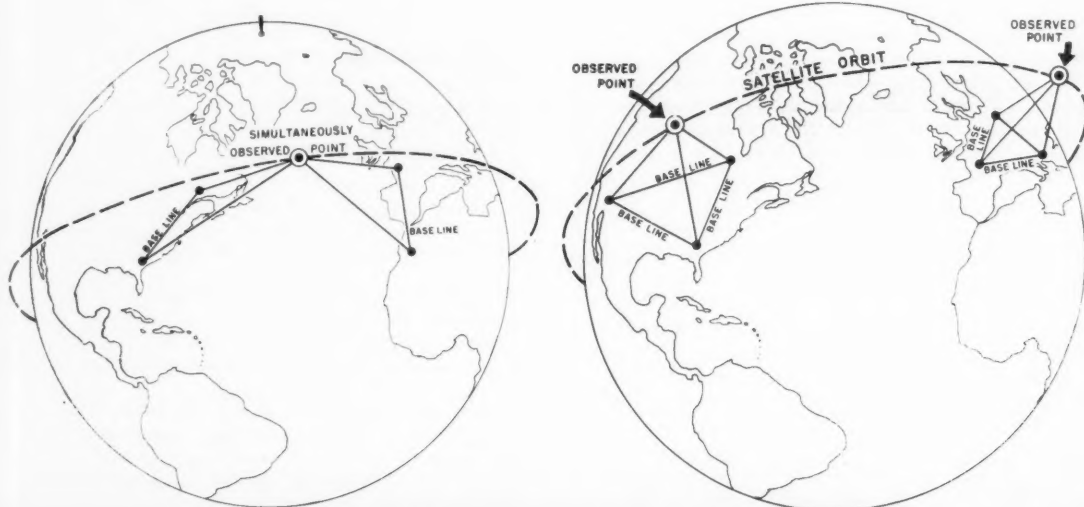


Fig. 1 Orbital and intervisible modes of geodetic satellite operation



Fig. 2 Schematic outline of primary world triangulation

ments $R\Delta\phi_0$, $R\Delta\lambda_0$, ΔR .⁴ In terms of geocentric spherical coordinates (ρ , ψ , λ), each of the displacement components $\Delta\rho$, $\rho\Delta\psi$, $\rho\Delta\lambda$ probably does not exceed 500 ft for the major datums such as North American and European. In addition to these systematic errors with respect to the geocentric coordinate system, major triangulation systems are subject to random errors reputedly of the order of one part to one half part in 10^6 in length for distances of 125 to 1000 miles. See Simmons(2)⁵.

For stations not referred to one of the major datums, the uncertainty in geocentric position increases rapidly to the limiting cases of isolated islands, the Antarctic, or other remote areas where an astronomic direction alone is used to estimate position in the geocentric system. Errors in ψ and λ of perhaps 2000 ft and of 500-1000 ft in ρ (if a level survey is not available) are possible under such circumstances.

Criteria for Optimum Geodetic Use of an Artificial Earth Satellite

Fig. 2 shows the approximate first-order triangulation coverage in the world today. A most significant feature to a group such as this one is the relatively small proportion of the Earth's surface covered by first order triangulation systems where it is presently possible to estimate the geocentric posi-

tion even approximately. Yet, all orbital motion of satellites and rockets is inherently referenced to the geocentric system rather than to one of the local ellipsoidal systems. Furthermore, most requirements for very precise positions on the surface of Earth now are associated with functions of the surface of Earth. Consequently, the station locations must be in geocentric coordinates in order to be compatible. This kind of analysis leads to some simple but extremely important ground rules regarding satellite geodesy. First, the objective of any geodetic satellite program should be to obtain maximum accuracy of station positions in a geocentric coordinate system. It should be noted that the notion of positional accuracy itself should be consistent with geocentric coordinates. If σ_p^2 is the variance in length units of a station location p on the surface of Earth, then

$$\begin{aligned}\sigma_p^2 &= \sigma_x^2 + \sigma_y^2 + \sigma_z^2 && \text{(Cartesian geocentric)} \\ \sigma_p^2 &= \sigma_\rho^2 + \rho^2\sigma_\psi^2 + \rho^2\sigma_\lambda^2 && \text{(spherical geocentric)} \\ \sigma_p^2 &\approx \sigma_R^2 + R^2\sigma_\phi^2 + R^2\sigma_\lambda^2 && \text{(ellipsoidal geocentric or "geographic")}\end{aligned}$$

This notion of positional error differs substantially from that commonly used in geodesy which only involves the relative error in distance or angle between two points usually referenced to the same datum. Consequently, our first criteria is simply to minimize σ_p^2 .

Next, we must make some statement about the number of stations to be located (or improved) in geocentric coordinates. Considering the number of stations independently of their distribution, it is obvious we wish to have a maximum number of stations. But the distribution of stations is quite important. One rough guide is that distribution should some-

⁴ Rotational terms in the transformation equations can be expected to be generally of less significance than the random surveying errors because of the high degree of parallelism of the geodetic and geocentric axes.

⁵ Numbers in parentheses indicate References at end of paper.

what parallel present density of triangulation or some other crude measure of eventual need for accurate positional information. However, a number of other minimal conditions should be considered also. All datums, however small or inaccurate, should have at least one good observation station occupied therein—better, at least three separated sites to provide for a complete adjustment and, ideally, several tens of such sites for redundancy. Isolated islands should be occupied at least once. All localities where precise tracking equipment of any kind is now located or is expected to be located should be occupied during the useful life of a geodetic satellite. Finally, at least one observation should be made in each 5 x 5 deg area of Earth's land surface and, depending on the availability of islands, over the ocean surface as well. Where feasible, the mesh size should be decreased to 1 x 1 deg. The primary value of the 5 x 5 deg and 1 x 1 deg stations is that the geocentric positions, along with the leveled heights of those stations, permit the geoid to be positioned in geocentric coordinates on a world-wide basis independently of, or in conjunction with, gravity and astrogeodetic data. The value of such geoid determination increases particularly as the observational accuracy approaches 100 ft or better.

Before summarizing the ground rules just outlined, an obvious simplification can be made regarding distribution and priority. There is no need, of course, to remeasure any significant fraction of the stations already tied into a datum in order to obtain geocentric positions of an accuracy not in excess of the internal accuracy of the datum. Three separated observations are sufficient to determine the three (3) components of a displacement vector such as to bring the local geodetic system into the geocentric one. Certainly, considerable redundancy is desirable in addition. But, it is clear that relatively few observation stations—e.g., 30 on a major datum—will improve the geocentric positions of thousands of control points already tied into that datum. Hence, the priority of about 30 such stations repeated perhaps ten times each is a maximum. Conversely, a much greater number of stations leading to observations of about the same accuracy as that of the datum itself probably will not be of great value because the basic inaccuracy of the triangulation itself will prevent indefinite increase of accuracy.⁶

Now let us summarize the criteria for the optimum geodetic use of a satellite:

- 1 Minimize σ_p ,² the absolute geocentric variance.
- 2 Maximize the total number of stations, but subject to the following conditions:
 - (a) Highest priority be given to about 300 observations from 30 distributed stations each on North American datum, European datum, Tokyo datum, and Indian datum.
 - (b) Second priority be given to other datums in South America, Asia, Australia, and New Zealand, and various other isolated datums.
 - (c) Third priority be given to all precise tracking station locations, existing or planned, not included in (a) or (b).
 - (d) Fourth priority be given to isolated island groups not included in above.
 - (e) Fifth priority be given to the 5 x 5 deg areas not included in above.
 - (f) Sixth priority be given to 1 x 1 deg areas not included in above and easily accessible.

It should be noted that since the number of stations required increases rapidly from (a) to (f), the final observational program must be based on a careful cost-time priority study if maximum effectiveness is to be made of the unavoidably costly field observation programs.

Observation Techniques

Now let us return to the problems of nominal system ac-

⁶ Except for the possibility of new, superaccurate baselines being set up.

curacy, observing station distribution, frequency of observations, and other questions involving observation parameters. Inasmuch as the major datums have highest priority for location of observing stations, and already are the most accurately located in geocentric coordinates, the criterion for the minimum improvement of σ_p by a geodetic satellite system is simply that it improve the average σ_p , $\bar{\sigma}_p$, of the major datums. An estimate⁷ of the very worst value of σ_p at present for the major datums might be 800 ft, but 500 ft is a more likely, but still pessimistic, estimate. Accordingly, the expected σ_p of an average observing station must be significantly less than 500 ft for a particular geodetic satellite system to be of any real interest. Just how much less is a problem of degree, but I would recommend the value of $\sigma_p = 200$ ft as the absolute minimum accuracy criterion for any satellite system which is designed primarily for geodetic purposes. This is about the largest value of σ_p which can reasonably be expected to result in statistically significant corrections to the presently inferred geocentric positions of observing stations located on major datums. If the expected σ_p is to be between 200 and 500 ft, the probability becomes considerable that such a geodetic satellite system would be *unable* to detect errors in the presently presumed geocentric positions of those observing stations. The upper limit to the desired accuracy of a geodetic satellite system has less immediate significance, of course, because the lower limit itself will require the use of only the most precise observation equipment presently available. It is worthwhile to note, however, that if a σ_p less than about 20 ft were desired, serious difficulties would arise with regard to existing baseline accuracy. Consequently, the accuracy requirement for a geodetic satellite system should be that the σ_p of the observing stations will be reduced to between 200 and 20 ft. These numbers correspond to about 10^{-5} and 10^{-6} Earth radii respectively.

The required system accuracy is thus $200 \text{ ft} > \sigma_p > 20 \text{ ft}$. The required accuracy of individual observations is not completely specified by σ_p because the orbit must be considered also in a complete error analysis. For the moment, however, it is perhaps reasonable to say that the systematic⁸ component of observation error E_s in a typical case certainly must be no larger than $\sigma_p/2$. If ten good observations are considered to be a typical number to be acquired in practice at a single site, then the random⁹ component E_R will be reduced by about $1/3$. Crudely then, we can conclude that, as a minimum condition

$$E_s + E_R/3 < \sigma_p$$

For the worst case, $\bar{\sigma}_p = 200$ ft, this admittedly crude approach suggests a maximum value of E_R of 300 ft in conjunction with a value of E_s of 100 ft. The design value of E_R in practice probably ought to be more on the order of 50 ft because (a) this entire analysis has been based on the worst case, and (b) much of the contribution to E_s cannot be easily reduced in practice since it is related to the station errors themselves.

The only completely proved observation system capable of meeting such an accuracy requirement for, say, a satellite at 1000 n mi slant range and arbitrary elevation angle is the ballistic camera system, of the type developed for precise missile tracking, used in conjunction with a powerful strobe light package in the satellite. Brown (3,4) describes the impressive data reduction and error analysis of ballistic

⁷ All estimates of $\bar{\sigma}_p$ must necessarily be based primarily on indirect methods since there are no continuous land bridges around the world.

⁸ "Systematic" error here refers to the component of total error which is highly correlated for that station from satellite pass to satellite pass. Systematic errors as defined here stem primarily from initial station errors themselves.

⁹ "Random" error refers to that component of the error which is uncorrelated from pass to pass. All observations on a single pass are assumed to be reduced to a single one by some reasonable method.

camera data as presently carried out at the Atlantic Missile Range. Errors in α' and δ' , the apparent right ascension and declination of a flashing light source, of the order of 2-4 sec can be expected with present systems. Much of this error is random, due to plate reading error and the effects of "seeing," and is reduced 2-4 fold by redundant images. Accordingly, an accuracy of 2 sec of arc (1:100,000) for two of the necessary three position coordinates of a satellite can reasonably be expected from a single ballistic camera system. The time of the observation is also acquired with the necessary accuracy by the ballistic camera system, and the systems are completely mobile. Several dozen suitable ballistic cameras are now in use or will be so shortly, including 36 in.-40 in. focal length aerial cameras lens systems which have been modified both by Air Force Cambridge Research Laboratories and Goddard Space Flight Center to perform in the ballistic camera mode with perhaps twice the accuracy mentioned above.

The Baker-Nunn satellite tracking cameras are probably capable of 2-4" accuracy also, but, to date, have not been used with adequate sources, time recording equipment, or time synchronization equipment to accumulate any substantial number of satellite observations of such precision. It seems reasonable to suppose that within two years or so the existing and planned Baker-Nunn tracking stations could provide observations of 2-4" accuracy and of error distribution similar to that of the ballistic camera systems. There will probably be about 18 well distributed stations in operation.

Finally, some astronomical telescopes appear to be capable of 0.5-1.0 sec accuracy, when used with a flashing light source, but such accuracy has not yet been demonstrated. As with the Baker-Nunn cameras, no provision exists at present for precise timing of the flash. Also, the relatively small field of astronomical telescopes restricts the number of redundant flashes recorded on a single plate, and thus cancels to some extent the benefits of the lower random error generated during measurement of plate coordinates. In summary, then, there are at present some stellar photography equipments capable of providing 2 sec accuracy of α' and δ' . Within a short time substantial numbers of both mobile and fixed stations capable of this accuracy on a routine basis will exist, as well as perhaps a few stations capable of 1 sec accuracy. The 2 sec accuracy figure corresponds to 60 ft at 1000 n mi and hence is roughly consistent at that range with the desired E_R .

There are, however, serious limitations associated with the use of optical positioning equipment with a geodetic satellite. First, the angular error corresponds to a linearly increasing amount of error in actual satellite position as the slant range is increased. Two seconds of arc at 5000 n mi slant range correspond to 300 ft—the upper limit for random error according to the previous analysis. Secondly, the accuracy and availability of observations begins to fall off below 20 deg elevation angle, rapidly below 10 deg. This loss of accuracy is mainly due to nonlinear atmospheric distortion coupled with the reduced number and inferior distribution of useful stellar images in the vicinity of the unknown image. Also, horizon visibility in all directions is rarely achieved at field observing sites. Thirdly, and most serious, is that optical observations are restricted to nontwilight night-time periods, and to periods of good to excellent visibility. Assuming an average of 45% useful darkness per day and 50% adequate visibility (rather optimistic for many places in the world), less than $\frac{1}{4}$ of the satellite passes theoretically visible at a typical station could be utilized. If simultaneous observations from two or more widely separated optical stations are desired, the likely number per month of successful intervisible optical observations becomes small indeed. Such considerations raise the specter of exorbitant ground station costs per useful observation. The exceedingly important notion of *efficiency* of a geodetic satellite system is thus encountered. It is clear that the cost per useful observation should be a minimum. In the case

of mobile stations, this means that the rate of collection of useful observations must be maximized. In practice, observations from permanent stations often will be available only at particular times due either to orbital characteristics or to conflicting demands on the station equipment (i.e., astronomical telescopes). Accordingly, it will be desirable to maximize those observations during a particular time interval. Thus, it is generally true that the efficiency of both mobile and permanent observing stations is increased by maximizing the rate of collection of observations.

For this reason, precision radar ranging techniques, as discussed in (5), for example, despite their presently inadequately demonstrated accuracy and reliability for satellite applications, are of paramount interest in any geodetic satellite program. The nearly all-weather capability of a radar ranging system leads to an inherently more efficient ground observing program than any optical system by a factor of 3 or 4—much more in some cases. The main present drawback to the use of precision radar ranging, either C W, or pulse, is that, despite manufacturer claims of accuracy, there appears to be little experimental data, which has been collected under controlled conditions, pertaining to ranging accuracy at microwave frequencies for a target above the ionosphere. Extrapolations of low altitude short-range test data strongly suggest that 50 ft or less range errors might be obtainable over a considerable variety of slant ranges, assuming sufficient transponder output power in the satellite.⁹ For this reason it seems reasonable to consider radar ranging observations to be of at least equal importance as optical for any truly geodetic satellite launched, say, two years or more hence. By such time, properly controlled tests of accuracy will have been carried out, and transponders and other satellite components can be expected to be demonstrably reliable for use in a geodetic satellite.

Let us summarize the error dependence of the stellar photography and radar ranging techniques for positioning a geodetic satellite:

1 Accuracy decreases with increasing slant range. For a fixed optical frequency or microwave power output from the satellite, the accuracy may fall off somewhat worse than linearly with increasing slant range for the radar case beyond some nominal short range, whereas optical accuracy can be considered to be more closely linear in range dependence.

2 Accuracy decreases with decreasing elevation angle, particularly below 10 deg.

These conclusions lead immediately to two more which are of great significance with regard to the choice of an optimum orbit:

1 Minimize the average slant range and the variability of slant ranges of the useful observations of all the stations.

2 Avoid low elevation angles; i.e., certainly below 10 deg.

Before proceeding on to a consideration of the choice of an orbit, it is useful to note the advantages of simultaneous radar range and stellar photographic observations. Inasmuch as the error in range lies along the range line whereas the angular errors are normal to it, simultaneously range and angle observations could be adjusted statistically to achieve an accuracy greater by a factor of 2 or more than the separate components of position. Similarly, if independent radial velocity measurements of geodetic accuracy were also acquired simultaneously, even greater control can be exercised on the

⁹ The situation regarding variable slant ranges with a fixed power output may be somewhat worse in the ranging case compared to the optical because the signal to noise ratio is apparently more marginal even under nominal circumstances. It is difficult to justify this assertion rigorously, but all the material available to this author regarding proposed geodetic accuracy of ranging systems has indicated that power available in the satellite, and hence power output from the satellite, is the limiting technical factor. Hence the signal to noise ratio will, in fact, be marginal on the long side of the optimum design range. In practice, then, accuracy may well fall off as badly or worse than in the optical case if the slant range is substantially increased above the design range.

accuracy of the estimated satellite positions. Such an adjustment could be particularly enhanced by taking into account certain known dynamical relationships between velocity and position as described by Murray and Lees (6). In all cases, however, the component accuracies must be of the same order of magnitude for there to be any significant benefit through simultaneous adjustment.

Orbit Considerations

We are now in a position to consider the optimum orbit for use in the intervisible method or in the orbital method, and to determine the relative merits of each. The two primary objectives of a geodetic satellite system, to minimize σ_p^2 and to maximize the number of stations subject to certain distribution conditions, generate four specific rules which dictate clearly the proper orbit regardless of the mode of observation. The rules are:

- (a) Maximize the rate of obtaining useful observations per station.
- (b) Minimize the average slant range of each useful observation.
- (c) Minimize the variability of these slant ranges.
- (d) Avoid low elevation angles, particularly below 10 deg.

The orbital method will be considered first. Two variations of this method may be used. In the short arc¹⁰ approach, two sets of observations of the satellite from disconnected observing sites (i.e., two discontinuous datums) separated in time by perhaps half a revolution or less are to be connected. The instantaneous orbital elements referenced to an osculating, perturbed ellipse are to be used as an accurate description of the exact path of the satellite between the two sets of observations. These instantaneous orbital elements could, for example, be obtained by smoothing the orbital elements of the satellite for, say, several days or weeks before and after the particular observation times in question. The path of the satellite connecting the stations would be obtained by numerical integration over the 15-60 min period in question. Corrections to station coordinates obtained in this manner should generally be as accurate as the observations themselves.

The interim or definitive orbit¹⁰ approach is simply, as the name suggests, to compute such an orbit carrying corrections to station coordinates as unknowns to be determined through minimization of the residuals. This method takes into account much more information in arriving at the station corrections and consequently potentially can supply corrections to an accuracy substantially greater than that of the observations themselves. Systematic observational errors, in the sense defined in this paper, represent the limit to the accuracy obtainable. Considering the orbital method in the light of rule (a) above, the interim or definitive orbit approach would be preferred because every observation of the satellite from every station makes its maximum contribution to the determination of corrections to the station coordinates. For example, a few good observations from Antarctica collected over several months could lead to corrected station coordinates there in an interim orbit even if no other stations observed the satellite during each of those particular revolu-

tions. In regard to (b) the consequence is clear: *minimize the average satellite height*, regardless of which variation of the orbital method is to be used.

The lower limit of the height is primarily determined by the accuracy of the dynamic model used to describe the satellite motion. The irregular departures of drag from the model become increasingly difficult to handle below about 300-400 statute miles height of perigee. In this regard, some unpublished computations carried out at the National Space Surveillance and Control Center (Formerly Project Space-track) are of interest. The satellite Vanguard I, with perigee height of about 400 miles and apogee height of about 2500 miles, was considered roughly comparable to a future geodetic satellite in regard to drag fluctuations. Mean values of \dot{P} , the rate of change of the period, have been tabulated (12). Typical values were selected and the equivalent displacements along the orbit computed. These displacements amount to only 3 to 4 m per revolution. As a check, the same satellite orbit was simulated on a large electronic computer by a numerical integration technique. The simulation was made first with no drag included, and then with a drag model based on the 1959 ARDC model atmosphere. The differences between the two simulations amounted to 3-10 m per revolution along the orbit and constitute an excellent check of the inductive approach. Accordingly, it is quite probable that the total drag effect on the Vanguard I satellite is not greater than 10 m per revolution and probably less than 5. Furthermore, that satellite has a very poor ratio of mass to surface area. A geodetic satellite would have a ratio 5-10 times as large. Finally, even a crude correction for *average* drag would reduce the error displacements substantially. It appears, therefore, safe to presume that for an orbit similar to that of Vanguard I the error would be less than one meter per revolution in the short arc method, and virtually negligible in the interim orbit approach. Even if unusual variations in the drag were to be encountered in a manner somehow neglected by the preceding analysis, these variations would have to be several orders of magnitude larger than the average rate of observations of the satellite by all stations, in order to introduce serious difficulties.

Thus a 400-mi perigee height appears to be quite safe for use in the orbital method; certainly there can be no question about a 500-mi perigee height. Perigee heights much below 400 miles require more detailed analysis, but probably would not be of interest even if the drag could be handled because the number of observations per revolution begins to be too few for efficient utilization of the ground stations. Accordingly, an optimum perigee height of 400-500 miles is indicated if the satellite is to be used in the orbital mode.

Rule (c), minimize the variability of the slant ranges, indicates that the orbit should be of low eccentricity. An apogee height of 500-700 statute miles might thus be indicated for optimum application of the orbital method. Rule (d), avoid low elevation angles, is equivalent to (b) for the orbital method.

The optimum orbit for the intervisible method is radically different. The primary requirement is that the satellite should be high enough during part of its orbit, so that the longer ties of 3000 to 4000 statute miles in length can be accomplished and yet also be low enough at other times to provide favorable geometry for shorter ties. An eccentric orbit is thus indicated. Fig. 3 is an approximate diagram¹¹ to demonstrate the problem. Using 20 deg as the desired elevation angle, for both atmospheric and geometrical reasons, it can be seen that the required satellite heights for ties of 2000, 3000, and 4000 mile distances are about 550, 850, and 1700 miles respectively. However, these heights only refer to a crossing at the exact center of tie. If, for example, one third of the passes over that tie are to be potentially useful and the elevation angle at one end is held fixed at 20 deg, whereas

¹⁰ The short period effects of the longitudinal variations in Earth's gravity field are not so well known at present for either method to be unambiguously applied. However, the work of O'Keefe and Batchlor (7), Musen (8), Izsak (9), Kaula (10), and Lees and Murray (11) indicate that some tesseral harmonics can be determined from satellite observations produced by the more accurate tracking systems already in use. Accordingly it is reasonable to suppose that by the time geodetic satellite data are actually to be reduced, the effects of the tesseral harmonics will be adequately known and can be taken into account. In any case, the geodetic satellite observations themselves will yield tesseral harmonic data.

¹¹ Based on Ref. (11).

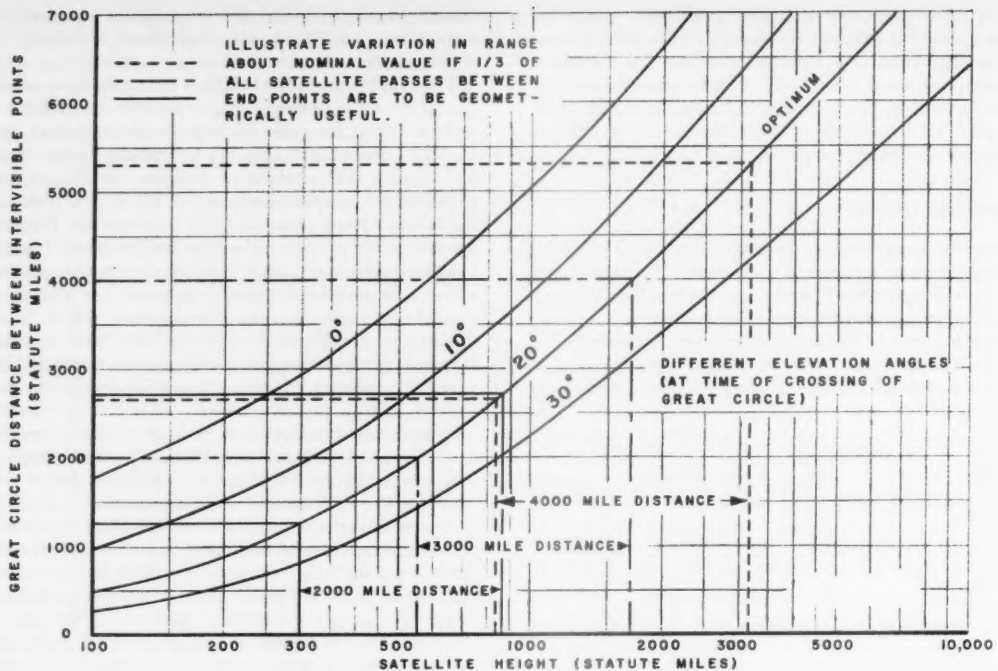


Fig. 3 Satellite altitude vs. range and elevation angle

that at the other end is allowed to vary from about 5 to 30 deg, then the required satellite heights vary from about 300 to 3200 statute miles. Accordingly, a perigee height of 300-400 statute miles and an apogee height of 2500-3500 statute miles might be suggested on this basis. Also, it is quite important that the perigee precess rapidly in the orbital plane in order that the full spectrum of satellite heights be available often throughout the world.

How does this optimum orbit for the intervisible method compare with the orbital method according to the four rules derived earlier? Very poorly, indeed! In the orbital method, the slant ranges would vary from 400-500 statute miles minimum (overhead pass at perigee) to 1500 statute miles maximum (700-mi apogee height and 20 deg elevation angle). In the intervisible method, on the other hand, the minimum slant range would be 700 statute miles (20 deg elevation angle with 300-mi perigee height) to a maximum distance in excess of 5000 statute miles (20 deg elevation angle with 3000 mi apogee height). The average slant range in the orbital case would be about 700 statute miles (55 deg elevation angle and 600 mi average satellite height). The average slant range in the intervisible case would be about 2800 miles (20 deg elevation angle, 1700-mi average satellite height), and certainly no less than 2100 mi (20 deg elevation angle, 1200 statute miles average satellite height for selected observations). The average slant ranges differ, therefore, by a factor of 3 to 4. Observational accuracy would be poorer in the intervisible case, for identical ground and satellite equipment, by a factor of at least 3 to 4. It is clear then, that the intervisible method suffers greatly by comparison with the desired minimization of slant ranges and variability of slant ranges (rules (b) and (c)).

In regard to maximization of the rate of collection of observations at each station, the situation is even worse. In the orbital method, every pass of the satellite near a station presents an opportunity for a useful observation. In the intervisible method, on the other hand, only those passes

which are simultaneously visible at, at least, two stations on each of, at least, two isolated datums can ever be used. On the basis of the geometry alone, this constraint reduces enormously the number of useful observations per station per unit of time. For example, if one-third of all passes across the great circle ties to be measured are useful when the satellite is within certain altitude limits, and the satellite is within those limits, say, one-half of the time (certainly a generous estimate), and, further, that 60% of these passes cross any desired great circle for a tie from that station to others, only 10% of the total number of satellite passes near a station could be useful. What is worse, any intercontinental intervisible method in the near future would probably have to use optical positioning only.¹² Consequently, even if the station under consideration is clear, the probability that three other separated stations are also clear (and in the dark night-time) is probably something worse than $(0.5)^3$; i.e., about 0.10. Therefore, the percentage of useful observations per station per unit in the intervisible method could be expected to be of the order of 1% of those in the orbital method.

Finally, the intervisible method has an even more severe limitation. It does not directly furnish corrections to geocentric coordinates at all! Only relative positions in geodetic coordinates are obtained. Geocentric coordinates can only be obtained by completing great circle links around the world, and coordinates obtained this way will be influenced by the weakest intervisible tie in the world-wide network.

It must be completely clear, then, that the intervisible method is so inferior in accuracy and efficiency as compared to the orbital method that it cannot be seriously considered as a means of improving any substantial number of station positions. The question then arises, as to what purpose it might serve. Its only real advantage is that no reliance is

¹² Because of the poor error propagation characteristic of oblique trilateration schemes.

placed on an orbital theory.¹³ Accordingly, it can be viewed as a complimentary check of the orbital method—although errors in orbital theory would have to be rather large in order to be detected by the intervisible method. Also, such intervisible ties can be accomplished much more efficiently and at least as accurately through the use of the rocket-flare triangulation technique.

The most pertinent conclusions to be drawn from the preceding analysis are:

1 The orbital and intervisible methods are incompatible with any single satellite. A clear choice as to the method must be made. For instance, an orbit of an average height of 1200 statute miles is too low to be of real use in the intervisible method, yet is twice the average height of a completely safe orbit for the orbital method. The positions thus determined with the higher satellite would be *two to four times* less accurate than with the lower satellite, the ground and satellite equipment being identical in both cases.

2 The intervisible method is substantially less accurate and enormously less efficient than the orbital method. The only apparent use for a satellite in the intervisible method would be as a check of the orbital method.¹⁴

Conclusions

1 The objective of a geodetic satellite program should simply be to minimize the geocentric variance σ_p^2 of a maximum number of stations. The priority of stations should be: (a) 30 or so stations on each of the major datums, (b) similarly for minor datums, (c) present or anticipated tracking stations, (d) isolated islands, (e) 5×5 deg squares, and (f) 1×1 deg squares as feasible.

2 The objective value of σ_p should be $200 \text{ ft} > \sigma_p > 20 \text{ ft}$; i.e., between 10^{-5} and 10^{-6} Earth radii.

3 Ballistic camera systems can provide the required accuracy now. Other optical systems and precise radar-ranging systems of adequate accuracy will be available soon. Doppler radial velocity measurements are of great value in conjunction with position measurements for a geodetic satellite.

4 Radar-ranging or, possibly, Doppler radio tracking should be the primary observation technique for a geodetic satellite, because of its all weather, day and night capability.

Optical measurements would be of additional value if the power and weight requirements of the optical package could be provided for.

5 The orbital method should be used with a satellite of 400–500 statute mi perigee height, 600–700 mi apogee height.

6 Any use of a satellite by the intervisible method should be justified, if at all, primarily as a check to the orbital method—not as a desirable way to use a satellite to improve geocentric station coordinates.

Acknowledgment

The author has benefited from valuable discussions of the geodetic satellite problem with many persons, particularly C. Geoffrey Hilton, Donald A. Rice, Alan B. Lees, and Owen W. Williams. These people, of course, are not responsible for any of the opinions expressed here.

References

- 1 Veis, G., "Geodetic Uses of Artificial Satellites," *Smithsonian Contr. Astrophys.*, vol. 3, no. 9, 1960.
- 2 Simmons, L., "How Accurate is First-Order Surveying?," *J. Coast & Geodetic Surv.*, no. 3, April 1950.
- 3 Brown, D. C., "A Treatment of Analytical Photogrammetry," RCA Data Reduction Technical Rep. no. 39, AFMTC, Patrick AFB, Fla., Aug. 1957.
- 4 Brown, D. C., "Photogrammetric Flare Triangulation," RCA Data Reduction Technical Rep. no. 46, AFMTC, Patrick AFB, Fla., Dec. 1958.
- 5 "Radar Ranging Equipment for a Geodetic Satellite," Systems Laboratory Division of Electronic Specialties Co., Los Angeles, Calif., May 3, 1960.
- 6 Murray, B. C. and Lees, A. B., "An Invariant Property of Satellite Motion in a Dissipative Medium," Proceedings of 7th Annual Meeting of the American Astronautical Society, Dallas, Tex., Jan. 1961.
- 7 O'Keefe, J. A. and Batchlor, C. D., "Perturbations of a Close Satellite by the Equatorial Ellipticity of the Earth," *Astron. J.*, vol. 62, 1957, pp. 183–185.
- 8 Musen, P., "On the Motion of a Satellite in an Asymmetrical Gravitational Field," *J. Geoph. Res.*, vol. 65, 1960, pp. 2783–2792.
- 9 Izsak, I. G., "A Determination of the Earth's Equator from the Motion of Two Satellites," *Smith. Inst. Astrophys. Obs., Research in Space Science Series, Special Rep. no. 56*, 1961, pp. 11–24.
- 10 Kaula, W. M., "Estimation of Longitudinal Variations in the Earth's Gravitational Field from Satellite Observations," Proceedings of 7th Annual Meeting of the American Astronautical Society, Dallas, Tex., Jan. 1961.
- 11 Lees, A. B. and Murray, B. C., "Special Geophysical Significance of Precise Satellite Velocity Observations," Abstract to appear in *J. Geoph. Res.*, July 1961.
- 12 Breggs, R. E., "A Table of Times of Perigee Passage for Satellite 1958 B₂," *Smith. Inst. Astrophys. Obs., Research in Space Science Series, Special Rep. no. 30*, Nov. 1959.
- 13 "The Geodetic Use of Artificial Earth Satellites," Parts I and II, Systems Laboratory Corp., Los Angeles, Calif., Aug. 15, 1959.
- 14 "Time Service Notice No. 7," U. S. Naval Observatory, Washington, D. C., June 19, 1959.
- 15 Whitney, C. A. and Veis, G., "A Flashing Satellite for Geodetic Studies," *Smith Inst. Astrophys. Obs., Research in Space Science Series, Special Rep. no. 19*, Dec. 1958.

¹³ Because of the development of VLF time synchronization techniques (14), the necessary clock accuracy at observing stations for the orbital method can be supplied on a routine basis.

¹⁴ For a kinder appraisal of the intervisible method see (15).

Relating Geodetic Latitude and Altitude to Geocentric Latitude and Radius Vector

E. W. PURCELL¹ and
W. B. COWAN²

Aeronutronic Div., Ford Motor Co.
Newport Beach, Calif.

The purpose of this paper is to develop a simple method of computing the geocentric latitude and the radius vector of a vehicle above an oblate Earth from its altitude and geodetic latitude. In addition, a solution to the important converse problem is also shown. An exact solution for an unfavorable location is given, confirming that the errors introduced by the recommended approximations are negligible. Although solutions have appeared previously, the following results are simpler, have not adversely compromised the accuracy, and are computed from other quantities which are themselves useful.

Preliminary Results

LET a AND b stand for the equatorial and polar radii respectively of the spheroid and let P denote a point with a geodetic latitude of μ and altitude h . (See Fig. 1.) From P , a normal to the ellipse at Q intercepts the semi-major axis OA at L making the angle PLA equal to μ and PQ equal to h . Let $OQ = R$, $OP = R$, $QL = C$, $OL = d$, angle $POA = \mu'$, angle $QOA = \mu''$, angle $OPL = \xi$, and angle $OQL = \delta$; O is the center of the ellipse.

The eccentricity e of the ellipse is given by

$$e^2 = 1 - \frac{b^2}{a^2} \quad [1]$$

It is clear that

$$\delta = \mu - \mu'' \quad [2]$$

$$\xi = \mu - \mu' \quad [3]$$

To derive an expression for δ , which is the difference between the geodetic and geocentric latitudes for points on the surface of the spheroid, we write the equation of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad [4]$$

Now

$$\frac{dy}{dx} = -\frac{xb^2}{ya^2} = -\frac{b^2}{a^2 \tan \mu''} = -\frac{1}{\tan \mu}$$

whence

$$a^2 \tan \mu'' = b^2 \tan \mu \quad [5]$$

which can be used in conjunction with Eq. 2 to find δ . However, since δ is small, a more suitable equation is found after some reduction from

$$\tan \delta = \tan (\mu - \mu'') = \frac{e^2 \sin \mu \cos \mu}{1 - e^2 \sin^2 \mu} \quad [6]$$

Using Eq. 5, one can eliminate μ to get

$$\tan \delta = \frac{e^2 \sin \mu'' \cos \mu''}{1 - e^2 \cos^2 \mu''} \quad [7]$$

Since δ is at most 11.5 min, $\tan \delta$ can be replaced by δ with an error of at most 1.3×10^{-6} radians which we neglect. We then have

$$\delta = \frac{e^2 \sin \mu \cos \mu}{1 - e^2 \sin^2 \mu} = \frac{e^2 \sin \mu'' \cos \mu''}{1 - e^2 \cos^2 \mu''} \quad [8]$$

These equations in conjunction with Eq. 2 are the basic equations relating geodetic and geocentric latitudes for points on the surface of Earth.

From $x = R_e \cos \mu''$ and $y = R_e \sin \mu''$ and Eq. 4

$$R_e = \frac{b}{\sqrt{1 - e^2 \cos^2 \mu''}} \quad [9]$$

Applying the law of sines to triangle OQL yields

$$d = \frac{R_e \sin \delta}{\sin \mu} \quad [10]$$

and

$$c = \frac{R_e \sin \mu''}{\sin \mu} \quad [11]$$

The law of cosines applied to triangle OPL now yields

$$R = \sqrt{d^2 + (c + h)^2 + 2d(c + h) \cos \mu} \quad [12]$$

Again the law of sines yields

$$\xi = \sin^{-1} \left(\frac{R_e}{R} \sin \delta \right) \quad [13]$$

which, since δ and ξ are always small, can be approximated by

$$\xi = \frac{R_e}{R} \delta \quad [14]$$

The error introduced by this approximation can be estimated by using series expansions for $\sin \delta$ and $\sin^{-1} x$. We then get, after dropping terms in δ of order higher than

Received Jan. 16, 1961.

¹ Research Engineer, Flight Analysis Dept.

² Engineer, Flight Analysis Dept. Member ARS.

$$e^2 = 1 - \frac{b^2}{a^2} \quad \text{Eq. 1}$$

$$\delta = \frac{e^2 \sin \mu \cos \mu}{1 - e^2 \sin^2 \mu} \quad \text{Eq. 8}$$

$$\mu'' = \mu - \delta, \quad \text{Eq. 2}$$

$$R_e = b \left[1 + \frac{e^2}{2} \cos^2 \mu'' + \frac{3}{8} e^4 \cos^4 \mu'' \right] \quad \text{Eq. 28}$$

$$\xi = \frac{R_e \delta}{R_e + h} \quad \text{Eq. 27}$$

$$R = R_e + h \left(1 - \frac{\delta \xi}{2} \right) \quad \text{Eq. 26}$$

$$\mu' = \mu - \xi \quad \text{Eq. 3}$$

Problem 2—The Converse Problem

Given the radius vector R of a vehicle and its geocentric latitude μ' , find the altitude and geodetic latitude of this vehicle.

The solution is given by the following equations

$$e^2 = 1 - \frac{b^2}{a^2} \quad \text{Eq. 1}$$

$$\delta' = \frac{e^2 \sin \mu' \cos \mu'}{1 - e^2 \cos^2 \mu'} \quad \text{Eq. 16}$$

$$R_e' = b \left[1 + \frac{e^2}{2} \cos^2 \mu' + \frac{3}{8} e^4 \cos^4 \mu' \right] \quad \text{Eq. 29}$$

$$\xi = \frac{R_e' \delta'}{R} \quad \text{Eq. 24}$$

$$h = (R - R_e') \left(1 - \frac{\delta' \xi}{2} \right) \quad \text{Eq. 23}$$

$$\mu = \mu' + \xi \quad \text{Eq. 3}$$

An Example

As an illustration, we have chosen the following example

$$a = 20,925,673 \text{ ft}$$

$$b = 20,855,523 \text{ ft}$$

$$h = 36,000,000 \text{ ft}$$

$$\mu = 0.7853981635 \text{ radians (45 deg)}$$

The following values were computed using exact equations throughout and carrying the results to ten significant figures at each step.

$$\delta = 0.003357947848 \text{ radian} \quad R_e = 20,890,745.22 \text{ ft}$$

$$\xi = 0.001233070051 \text{ radian} \quad d = 99,207.3144 \text{ ft}$$

$$\mu'' = 0.7820402157 \text{ radian} \quad c = 20,820,477.53 \text{ ft}$$

$$\mu' = 0.7841650934 \text{ radian} \quad R = 56,890,670.94 \text{ ft}$$

$$\delta' = 0.003357965446 \text{ radian} \quad R_e' = 20,890,596.16 \text{ ft}$$

Using Eqs. 1, 8, 2, 28, 27, 26 and 3, we get the following solutions to the problem of the first type

$$\xi = 0.001233070437 \text{ radian}$$

$$\mu' = 0.7841650931 \text{ radian}$$

$$R = 56,890,670.44 \text{ ft}$$

The errors are thus: 3×10^{-10} radian in μ' and 0.50 ft in R .

Using Eqs. 1, 16, 29, 24, 23 and 3, we get the following solution to the problem of the second type

$$\xi = 0.001233069705 \text{ radian}$$

$$\mu = 0.7853981631 \text{ radian}$$

$$h = 36,000,000.49 \text{ ft}$$

The errors are seen to be 4×10^{-10} radians in μ and 0.49 ft in h . These errors are considered to be negligible.

Conclusion

Although many of the above equations are approximations, the results in general are very good and will usually show relative errors less than those indicated by the above example. The reader will readily find from the above results several still simpler (but less accurate) formulas which may well be adequate for his applications.

References

- 1 Berger, W. J. and Ricupito, J. R., "Prediction Theory of Missile and Satellite Orbits," ARS JOURNAL, vol. 29, no. 6, June 1959, pp. 428-432.
- 2 Berger, W. J. and Ricupito, J. R., "Geodetic Latitude and Altitude of a Satellite," ARS JOURNAL, vol. 30, no. 9, Sept. 1960, pp. 901-902.
- 3 American Ephemeris and Nautical Almanac, for the year 1961, U. S. Government Printing Office, Washington, D. C., 1959, pp. 457 and 489.
- 4 Albright, N. W., "The Transformation of Earth-Referenced Data to Inertial Coordinate Systems," Progress Rep. no. 30-5, Jet Propulsion Lab., Pasadena, Calif., 1959.

Rocket Boost Vehicle Mission Optimizations¹

R. T. STANCIL² and
L. J. KULAKOWSKI³

Convair-Fort Worth
A Div. of General Dynamics Corp.
Fort Worth, Tex.

A complete analysis of rocket boost trajectory optimizations is presented to illustrate some of the advantages resulting from use of the indirect methods of the calculus of variations. The method devised is intended for boost vehicle evaluations during design studies. Features of the method include provisions to account for the effects of planetary rotation, aerodynamic forces, variable atmosphere and time-variant thrust. Vehicle design or physiological limitations are satisfied by employing inequality constraints involving the angle of attack. Special end conditions are developed for direct calculation of extremals for most rocket boost missions of interest. Computational techniques employed to obtain solutions on a high speed digital computer are discussed, with particular emphasis on convergence of the iteration to find the desired extremal. Numerical examples are shown to illustrate the character of extremals for satellite launchings and the minimal sensitivity of extremals to guidance and control errors.

DURING the past decade an extensive effort has been devoted to the optimization of aircraft and missile flight paths by employing the indirect methods of the calculus of variations, dynamic programming and gradient methods. Of these techniques, the indirect methods of the calculus of variations have been most popular, as indicated in the elegant surveys by Miele (1 and 2)⁴ and Breakwell (3). Kelley (4) recently indicated the computational advantages of gradient methods over indirect methods. The objective of this paper is to demonstrate a practical technique for solving indirect variational problems by overcoming some of the computational disadvantages that have plagued previous researchers.

A fairly formidable problem, involving the optimization of most rocket boost vehicle missions of interest, was selected as a framework for the demonstration. Simulations employed in this problem are sufficiently realistic to serve as a basis for making sensible engineering decisions in vehicle design studies. Solutions require nominal amounts of high speed digital computer time.

The present analysis represents an extension of a method reported previously (5). Principal extensions include incorporation of rotating Earth effects and development of special end conditions that permit the direct calculation of extremals for specific missions. A discussion of techniques providing improved convergence in the iteration to satisfy separated boundary conditions is included. One numerical example presented compares extremals for satellite launchings by two rockets having the same total impulse, with different burning times. The feasibility of simulating an extremal by practical guidance and control techniques is shown in another numerical example.

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ This work was supported under Convair-sponsored research programs.

² Project Aerodynamics Engineer, Aero Sciences Group.

³ Design Specialist, Aero Sciences Group. Member ARS.

⁴ Numbers in parentheses indicate References at end of paper.

Equations of Motion

The stated objective of this analysis is to develop a procedure useful for comparative vehicle evaluations in the design study phase. To satisfy this objective and, simultaneously, achieve reasonable computation times, certain facets of the analysis are expressly tailored to satisfy the requirements of design evaluations. For example, the equations of motion used represent the motion of a point mass in the vicinity of a spherical, rotating planet. However, since only the gross effects of rotation are sought, it is assumed that the rocket moves in a great circle plane fixed in inertial space. With this approximation, the time history of the trajectory of a high thrust, short duration (on the order of 15 min) booster will be quite accurate, with the principal effect being a small discrepancy in initial azimuth to provide the desired inclination at burnout. The approximation is justified by reduction in computation time since two additional Lagrangian multipliers, required in a three-dimensional representation, are eliminated.

Orientation of the rocket relative to the planet and the force, velocity, and acceleration vectors that exist during flight are shown in Fig. 1. Equations of motion used employ natural coordinates directed along and normal to the flight path

$$m\dot{v} = T \cos \alpha - D - gm \sin \gamma \quad [1]$$

$$m\dot{v} \left(\dot{\gamma} - \frac{v}{R+h} \cos \gamma - 2\omega \cos i \right) = T \sin \alpha + L - gm \cos \gamma \quad [2]$$

In addition, the kinematic relationship

$$\dot{h} = v \sin \gamma \quad [3]$$

is used. The Coriolis acceleration component ($2\omega v \cos i$) in the plane of motion is included in Eq. 2. Transverse Coriolis accelerations were excluded. Centrifugal terms of negligible

magnitude have been deleted from the motion equations. Consideration of thrust vector gimbaling was also excluded, to eliminate an additional dependent variable. Hence, the variable α represents both the angle of attack and the angle between the thrust and velocity vectors.

In performing the analysis it is assumed that the mass, thrust, lift, drag, and acceleration of gravity are known at all points of the trajectory and have the following functional dependence: $m = m(t)$, $T = T(h, t)$, $L = L(v, h, \alpha)$, $D = D(v, h, \alpha)$, $g = g(R + h)$. Lag effects in the action of the aerodynamic forces are excluded. The aerodynamic force coefficients may be discontinuous at staging points.

Basic Variational Problem

The basic variational problem considered is the same as that discussed previously (5). For the sake of completeness and the convenience of restating the problem in the Bolza form, the following brief summary is included.

Euler-Lagrange Equations

Eqs. 1, 2, and 3, written in the forms

$$\begin{aligned}\varphi_1 &= \dot{v} - F(v, \gamma, h, \alpha, t) = 0 \\ \varphi_2 &= \dot{\gamma} - G(v, \gamma, h, \alpha, t) = 0 \\ \varphi_3 &= \dot{h} - v \sin \gamma = 0 \\ &\dots\dots\dots [4]\end{aligned}$$

serve as auxiliary conditions on the variational problem, permitting arbitrary variations in the dependent variables v , γ , h , and α . In addition, the inequality constraint

$$\varphi_4 = L_m - l_m(v, \gamma, h, \alpha) \geq 0 \quad m = 1, 2, 3, 4 \quad [5]$$

is introduced to satisfy vehicle design and physiological limitations. The relationships

$$\begin{aligned}l_1 &= -\frac{T}{m} \sin \beta + \frac{L}{m} \cos(\alpha + \beta) + \frac{D}{m} \sin(\alpha + \beta) \\ l_2 &= q|\alpha| \quad l_3 = |\alpha| \quad l_4 = T_{sw}(v, h, \alpha)\end{aligned}$$

are the control functions employed to satisfy acceleration, loading, stability, and temperature requirements, respectively.

For the reasons given in (5), the basic variational problem involves maximizing the burnout velocity. Eqs. 4 and 5 are satisfied and provisions for including an optimal coasting stage are included. In terms of a Bolza-type problem (6), it is

$$\begin{aligned}\delta I &= \left[\left(f - \sum_{k=1}^4 \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \right) \delta t \right]_{t_0}^{t^*} + \left[\sum_{k=1}^4 \frac{\partial f}{\partial \dot{x}_k} \delta x_k \right]_{t_0}^{t^*} + \left[\left(f - \sum_{k=1}^4 \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \right) \delta \tau \right]_0^{t_f - t^*} + \left[\sum_{k=1}^4 \frac{\partial f}{\partial \dot{x}_k} \delta x_k \right]_0^{t_f - t^*} \\ &\quad + \sum_{k=1}^4 \left[\int_{t_0}^{t^*} \left(\frac{\partial f}{\partial x_k} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}_k} \right) (\delta x_k - \dot{x}_k \delta t) dt + \int_0^{t_f - t^*} \left(\frac{\partial f}{\partial x_k} - \frac{d}{d\tau} \frac{\partial f}{\partial \dot{x}_k} \right) (\delta x_k - \dot{x}_k \delta \tau) d\tau \right] + \delta x_{t_f} = 0 \quad [7]\end{aligned}$$

required to find a stationary value of the integral

$$I = \int_{t_0}^{t^*} f(t, x_k, \dot{x}_k) dt + \int_0^{t_f - t^*} f(\tau, x_k, \dot{x}_k) d\tau + x_{t_f} \quad \tau = t - t^* \quad [6]$$

where the x_k , $k = 1, 2, 3, 4$ represent v , γ , h , and α , respectively, and

$$f(t, x_k, \dot{x}_k) = \sum_{i=1}^4 \lambda_i(t) \varphi_i(t, x_k, \dot{x}_k)$$

In Eq. 6, t^* represents the time at the end of the coasting period, while t_0 and t_f are the initial and final time for the

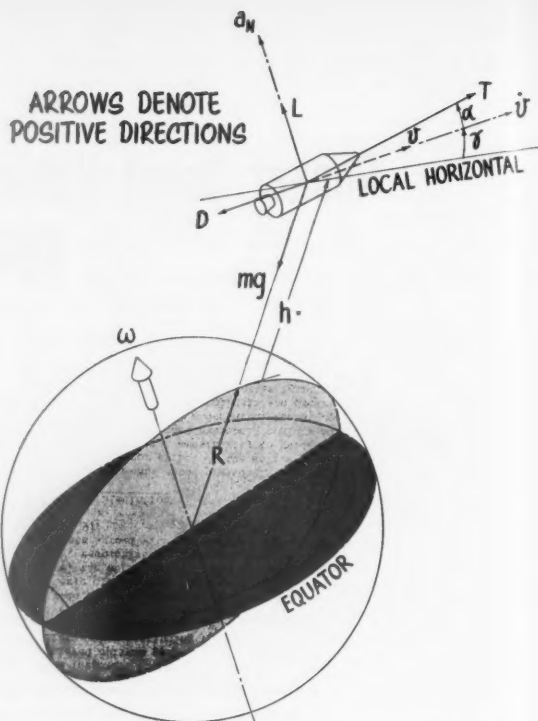


Fig. 1 Rocket orientation and vector system

problem, respectively. Note that the relationship $\lambda_4 \varphi_4 = 0$ must be satisfied in this formulation. A necessary condition for the existence of such a stationary value is that the first variation vanishes

$$\begin{aligned}\delta I &= [f(t) \delta(t)]_{t_0}^{t^*} + \int_{t_0}^{t^*} \delta f(t) dt + [f(\tau) \delta \tau]_0^{t_f - t^*} + \\ &\quad + \int_0^{t_f - t^*} \delta f(\tau) d\tau + \delta x_{t_f} = 0\end{aligned}$$

After substituting the appropriate expressions for δf and performing the usual integration by parts, the expression for the first variation becomes

$$\begin{aligned}\delta I &= \left[\left(f - \sum_{k=1}^4 \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \right) \delta t \right]_{t_0}^{t^*} + \left[\sum_{k=1}^4 \frac{\partial f}{\partial \dot{x}_k} \delta x_k \right]_{t_0}^{t^*} + \left[\left(f - \sum_{k=1}^4 \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \right) \delta \tau \right]_0^{t_f - t^*} + \left[\sum_{k=1}^4 \frac{\partial f}{\partial \dot{x}_k} \delta x_k \right]_0^{t_f - t^*} \\ &\quad + \sum_{k=1}^4 \left[\int_{t_0}^{t^*} \left(\frac{\partial f}{\partial x_k} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}_k} \right) (\delta x_k - \dot{x}_k \delta t) dt + \int_0^{t_f - t^*} \left(\frac{\partial f}{\partial x_k} - \frac{d}{d\tau} \frac{\partial f}{\partial \dot{x}_k} \right) (\delta x_k - \dot{x}_k \delta \tau) d\tau \right] + \delta x_{t_f} = 0 \quad [7]\end{aligned}$$

Since the variations δx_k are arbitrary, application of the basic lemma of the calculus of variations yields the Euler-Lagrange equations

$$\frac{\partial f}{\partial x_k} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}_k} = 0 \quad k = 1, 2, 3, 4 \quad [8]$$

or, in terms of the physical variables

$$\dot{\lambda}_1^* + \lambda_1^* \frac{\partial F}{\partial v} + \lambda_2^* \frac{\partial G}{\partial v} + \lambda_3^* \sin \gamma - \lambda_4^* \frac{\partial \varphi_4}{\partial v} = 0$$

$$\dot{\lambda}_2^* + \lambda_1^* \frac{\partial F}{\partial \gamma} + \lambda_2^* \frac{\partial G}{\partial \gamma} + \lambda_3^* v \cos \gamma - \lambda_4^* \frac{\partial \varphi_4}{\partial \gamma} = 0$$

$$\dot{\lambda}_3^* + \lambda_1^* \frac{\partial F}{\partial h} + \lambda_2^* \frac{\partial G}{\partial h} - \lambda_4^* \frac{\partial \varphi_4}{\partial h} = 0$$

$$\lambda_1^* \frac{\partial F}{\partial \alpha} + \lambda_2^* \frac{\partial G}{\partial \alpha} - \lambda_4^* \frac{\partial \varphi_4}{\partial \alpha} = 0 \quad [9]$$

where

$$\lambda_1^* = \frac{\lambda_i}{\lambda_{10}} \quad \dot{\lambda}_i^* = \frac{\dot{\lambda}_i}{\lambda_{10}}$$

In Eqs. 9 advantage was taken of the fact that the Euler-Lagrange equations are homogeneous in the multipliers λ_i . As noted by Cicala and Miele (7), solutions of the equations are independent of the initial values assigned to one of the multipliers. For the present application, λ_1 is the multiplier selected, and its initial values are arbitrary.

Corner Conditions

Prior to evaluating the transversality terms of Eq. 7, the behavior of the trajectory at staging points must be established. At such points the thrust and mass of the rocket are ordinarily discontinuous. Consequently, the derivatives \dot{v} and $\dot{\gamma}$ are also discontinuous across staging points, while the functions v , γ , h , and \dot{h} are continuous. The Erdmann-Weierstrass corner conditions supply the relationships

$$\left(\frac{\partial f}{\partial \dot{x}_i} \right)_- = \left(\frac{\partial f}{\partial \dot{x}_i} \right)_+ \quad [10]$$

where the minus and plus signs indicate functions evaluated just prior to and immediately after staging, respectively. Consequently, the multipliers λ_i , $i = 1, 2, 3$ are continuous across staging points.

Both the driving function α and the multiplier λ_4 may be discontinuous at staging points. It will be demonstrated later that discontinuities in α , while physically unrealistic, have little effect on solutions obtained for the complete trajectory.

Boundary Conditions

In all problems considered herein the velocity, path angle, and altitude are specified at launch; i.e., the initial conditions are

$$\text{when } t = t_0 \quad v = v_0, \gamma = \gamma_0, h = h_0 \quad [11]$$

For the basic problem of maximizing burnout velocity the final values of path angle Γ and altitude H are specified

$$\text{when } t = t_f \quad \gamma_f = \Gamma, h_f = H \quad [12]$$

To define the desired extremal, the transversality condition of Eq. 7 must also be satisfied. Since $f = 0$ on trajectories and the initial conditions are fixed by Eq. 11, the transversality condition becomes

$$\sum_{k=1}^4 \left\{ \left(\frac{\partial f}{\partial \dot{x}_k} \delta x_k - \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \delta t \right)_{t=t^*} + \left[\frac{\partial f}{\partial \dot{x}_k} \delta x_k - \dot{x}_k \frac{\partial f}{\partial \dot{x}_k} \delta \tau \right]_0^{t_f-t^*} \right\} + \delta x_{1f} = 0$$

After employing Eq. 10 and noting that the burning time of the stages following the optimum coast period is fixed, the transversality condition reduces to

$$-(\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{h})_{t=t^*} \delta t_c + (\lambda_{1f} + 1) \delta v_f + \lambda_{2f} \delta \gamma_f + \lambda_{3f} \delta h_f = 0 \quad [13]$$

where t_c is the total coasting time. On an extremal, $\delta v_f = 0$ and since the final path angle and altitude are specified by Eq. 12, Eq. 13 yields

$$(\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{h})_{t=t^*} = 0 \quad [14]$$

for arbitrary variations in coasting time.

To find the path yielding maximum burnout velocity, the initial conditions, Eq. 11, are combined with guessed values of λ_{20}^* , λ_{30}^* and t_c to initiate trajectory calculations involving the numerical integration of Eqs. 4, 9 and, if required, 5. The guessed parameters are then subjected to iteration until the boundary conditions, Eqs. 12 and 14, are satisfied. When an optimum coasting stage is not included in the problem ($t_c = 0$, $\delta t_c = 0$), Eq. 13 is automatically satisfied and the λ_{20}^* and λ_{30}^* that satisfy Eq. 12 yield the desired extremal. In the latter case every trajectory computed is an extremal for the burnout conditions achieved.

An interesting interpretation of Eq. 14 may be achieved by noting the effect on extremals of variations in the initial conditions. As indicated by Breakwell (3), consider two neighboring optimal trajectories which originate at the points (t_0, x_{k0}) and $(t_0 + \delta t_0, x_{k0} + \delta x_{k0})$ and satisfy the same constraints, Eq. 5, and end conditions, Eqs. 12 and 14. From Eq. 7

$$\lambda_{1f} \delta v_f = \lambda_{10} \delta v_0 + \lambda_{20} \delta \gamma_0 + \lambda_{30} \delta h_0 - \left(\sum_{k=1}^3 \lambda_{k0} \dot{x}_{k0} \right) \delta t_0$$

Now, upon selecting λ_{10} so that $\lambda_{1f} = -1$, it follows that

$$\frac{\partial v_f}{\partial v_0} = -\lambda_{10} \quad \frac{\partial v_f}{\partial \gamma_0} = -\lambda_{20} \quad \frac{\partial v_f}{\partial h_0} = -\lambda_{30} \quad \frac{\partial v_f}{\partial t_0} = \sum_{k=1}^3 \lambda_{k0} \dot{x}_{k0}$$

Furthermore, for unique trajectories

$$\frac{\partial v_f}{\partial v} = -\lambda_1 \quad \frac{\partial v_f}{\partial \gamma} = -\lambda_2 \quad \frac{\partial v_f}{\partial h} = -\lambda_3 \quad \frac{\partial v_f}{\partial t} = \sum_{k=1}^3 \lambda_k \dot{x}_k \quad [15]$$

at arbitrary times within the trajectory. The last of Eqs. 15 shows that Eq. 14 is equivalent to $\partial v_f / \partial t_c = 0$, thus confirming the correctness of Eq. 14 as the end condition for optimum coasting time. The significance of the derivative properties of the multipliers will be demonstrated later. When applied at burnout, these properties define pertinent derivatives of the "extremal surface," composed of the points (v_f, γ_f, h_f) , that permit generalization of the basic variational problem.

Applying a similar analysis to (unconstrained) extremals, it is possible to derive the relationship $\partial v_f / \partial \alpha = 0$ in addition to those of Eq. 15. This relationship verifies the minimal sensitivity of extremals to guidance and control errors.

Specific Mission Optimizations

The chosen formulation of the boost problem provides increased versatility with regard to the direct computation of trajectories that satisfy a variety of optimal mission requirements. Ultimate performance capabilities of a given booster/payload combination are represented by the "extremal surface" described previously. Optimal launch requirements for several types of missions form lines in the v, γ, h -space. Hence, by finding the intersection of a "mission line" with the "extremal surface" (assuming such an intersection exists), it is possible to write special end conditions for the extremal that optimizes the mission considered, while retaining the same Euler-Lagrange equations. Derivation of several of the special mission end conditions is made possible by the ability to formulate analytical expressions for pertinent derivatives of the extremal surface. Such end conditions replace those given by Eqs. 12. The optimum coasting time condition, Eq. 14, may be used in conjunction with the particular mission end conditions, if desired.

Hypersonic Glider Range

A simple extension of the basic variational problem involves finding the extremal that maximizes the equilibrium glide range of a hypersonic glide vehicle. At the initiation of gliding flight (booster burnout), $\dot{\gamma} = 0$ and Eq. 2 yields

$$mv \left(\frac{v}{R+h} \cos \gamma + 2\omega \cos i \right) + L - gm \cos \gamma = 0$$

After making some obvious substitutions, the burnout boundary conditions for a specified final path angle become

$$\gamma_f - \Gamma = 0$$

$$1 - \left\{ \frac{2}{p} \frac{W}{SC_L} \left[\frac{\cos \Gamma}{v^2} - \frac{\cos \Gamma}{g(R+h)} - \frac{2\omega}{gv} \cos i \right] \right\}_{t=t_f} = 0 \quad [16]$$

where the equilibrium glide parameter W/SC_L is prescribed for the given glider in free flight. The extremal in this mission option maximizes the energy available for equilibrium flight and, consequently, maximizes the glide range for any glider lift to drag ratio.

Extremal Surface Derivatives

In the subsequent mission optimizations the appropriate burnout boundary conditions are derived by employing the results of two-body orbits. Therefore, it is necessary to transform both the burnout conditions and the extremal surface derivatives to an inertial reference. The burnout altitude is not affected by the transformation. The inertial velocity \bar{v}_f (bars denote inertial-referenced quantities) is given by the vectorial sum of the planet-relative velocity v_f and the component of the planet's velocity of rotation

$$V = (R + h_f)\omega \cos i \quad [17]$$

Hence, applying the law of cosines

$$\bar{v}_f^2 = v_f^2 + V^2 + 2v_f V \cos \gamma_f \quad [18]$$

The law of sines is used to derive the expressions

$$\sin \tilde{\gamma}_f = \frac{v_f}{\bar{v}_f} \sin \gamma_f \quad \cos \tilde{\gamma}_f = \frac{1}{\bar{v}_f} (v_f \cos \gamma_f + V) \quad [19]$$

which determine the magnitude and quadrant of the inertial path angle.

Derivatives of the extremal surface may now be transformed. Consider two neighboring extremals that have the same initial conditions and constraints, but different final conditions. In this instance, Eq. 7 yields

$$\delta v_f = -\hat{\lambda}_{2f} \delta \gamma_f - \hat{\lambda}_{3f} \delta h_f \quad \hat{\lambda}_{1f} = \lambda_{1f} / \lambda_{1f} \quad [20]$$

Now, from Eq. 18

$$\bar{v}_f \delta \bar{v}_f = (v_f + V \cos \gamma_f) \delta v_f - v_f V \sin \gamma_f \delta \gamma_f + (V + v_f \cos \gamma_f) \omega \cos i \delta h_f \quad [21]$$

Combining Eqs. 20 and 21

$$\bar{v}_f \delta \bar{v}_f = -[\hat{\lambda}_{2f}(v_f + V \cos \gamma_f) + v_f V \sin \gamma_f] \delta \gamma_f + [(V + v_f \cos \gamma_f) \omega \cos i - \hat{\lambda}_{3f}(v_f + V \cos \gamma_f)] \delta h_f \quad [22]$$

From Eq. 20 and the first of Eqs. 19

$$\bar{v}_f \cos \tilde{\gamma}_f \delta \tilde{\gamma}_f + \sin \tilde{\gamma}_f \delta \bar{v}_f = (v_f \cos \gamma_f - \hat{\lambda}_{2f} \sin \gamma_f) \delta \gamma_f - \hat{\lambda}_{3f} \sin \gamma_f \delta h_f \quad [23]$$

After some algebraic manipulation involving the elimination of $\delta \gamma_f$ between Eqs. 22 and 23 and subsequent substitution of the second of Eqs. 19

$$-[\hat{\lambda}_{2f}(v_f + V \cos \gamma_f) + v_f V \sin \gamma_f] \delta \tilde{\gamma}_f +$$

$$[-\hat{\lambda}_{3f} v_f + (v_f \cos \gamma_f - \hat{\lambda}_{2f} \sin \gamma_f) \omega \cos i] \delta h_f = \frac{1}{\bar{v}_f} [v_f^2 + v_f V \cos \gamma_f - \hat{\lambda}_{2f} V \sin \gamma_f] \delta \bar{v}_f \quad [24]$$

Hence, the desired derivatives are

$$\frac{\partial \bar{v}_f}{\partial h_f} = \frac{\bar{v}_f [-\hat{\lambda}_{3f} v_f + (v_f \cos \gamma_f - \hat{\lambda}_{2f} \sin \gamma_f) \omega \cos i]}{v_f^2 + v_f V \cos \gamma_f - \hat{\lambda}_{2f} V \sin \gamma_f}$$

$$\frac{\partial \bar{v}_f}{\partial \tilde{\gamma}_f} = -\frac{\bar{v}_f [\hat{\lambda}_{2f}(v_f + V \cos \gamma_f) + v_f V \sin \gamma_f]}{v_f^2 + v_f V \cos \gamma_f - \hat{\lambda}_{2f} V \sin \gamma_f} \quad [25]$$

in the inertial reference frame. These derivatives are employed in most of the following mission optimizations.

Apogee Altitude

Maximizing the apogee altitude is of interest for sounding rockets and for precoasting boost trajectories in the launching of very high altitude satellites. From the two-body problem, the radius of apogee is

$$r_a = a(1 + e) \quad [26]$$

where, in terms of burnout conditions, the orbit eccentricity is

$$e = \sqrt{1 - \frac{\bar{v}_f^2 \cos^2 \tilde{\gamma}_f}{\bar{g}_f(R + h_f)} \left[2 - \frac{\bar{v}_f^2}{\bar{g}_f(R + h_f)} \right]} \quad [27]$$

and the mean distance is

$$a = \frac{\bar{g}_f(R + h_f)^2}{2\bar{g}_f(R + h_f) - \bar{v}_f^2} \quad [28]$$

where \bar{g}_f represents the central force field acceleration of gravity at burnout. In this analysis it is desired to maximize the apogee altitude for specified values of $\tilde{\gamma}_f$. This is achieved when the conditions

$$\frac{dr_a}{dh_f} = (1 + e) \left(\frac{\partial a}{\partial \bar{v}_f} \frac{\partial \bar{v}_f}{\partial h_f} + \frac{\partial a}{\partial h_f} \right) + a \left(\frac{\partial e}{\partial \bar{v}_f} \frac{\partial \bar{v}_f}{\partial h_f} + \frac{\partial e}{\partial h_f} \right) = 0$$

$$\tilde{\gamma}_f - \Gamma = 0$$

are satisfied. Taking the derivatives of Eqs. 27 and 28, these conditions become

$$2(1 + e) \left(1 + \frac{\bar{v}_f}{\bar{g}_f} \frac{\partial \bar{v}_f}{\partial h_f} \right) - \frac{\bar{v}_f \cos^2 \tilde{\gamma}_f}{ae \bar{g}_f} \left[2(R + h_f) \frac{\partial \bar{v}_f}{\partial h_f} + \bar{v}_f \right] \times \left[1 - \frac{\bar{v}_f^2}{\bar{g}_f(R + h_f)} \right] = 0 \quad \tilde{\gamma}_f - \Gamma = 0 \quad [29]$$

Analytical development of end boundary conditions for maximum apogee altitude without the restriction $\tilde{\gamma}_f = \Gamma$ may easily be achieved. However, an analysis of computed solutions indicated that convergence difficulties would preclude the satisfaction of such conditions. The ultimate apogee altitude extremal can be found by cross-plotting results derived from the adopted approach.

Ballistic Range

Ballistic range of the payload may readily be maximized by the proposed method. From the two-body problem, the ballistic range angle may be written in the form

$$\Phi = \frac{\sin \tilde{\gamma}_f}{\sin |\tilde{\gamma}_f|} \cos^{-1} \left\{ \frac{1}{e} \left[1 - \frac{\bar{v}_f^2 \cos^2 \tilde{\gamma}_f}{\bar{g}_f(R + h_f)} \right] \right\} + \cos^{-1} \times \left[\frac{1}{e} \left(1 - \frac{\bar{v}_f^2 \cos^2 \tilde{\gamma}_f}{\bar{g}_f R} \right) \right]$$

The ballistic missile "hit equation," given by Wheelon (8), is

$$\frac{R + h_f}{R} - \frac{\bar{g}_0 R^2 (1 - \cos \Phi)}{\bar{v}_f^2 (R + h_f) \cos^2 \tilde{\gamma}_f} - \frac{\cos (\tilde{\gamma}_f + \Phi)}{\cos \tilde{\gamma}_f} = 0 \quad [30]$$

where \bar{g}_0 is the acceleration of gravity at the planet's surface. Eq. 30 yields the error equation for ballistic paths

$$\left[1 + \frac{\bar{g}_f R(1 - \cos \Phi)}{\bar{v}_f^2 \cos^2 \gamma_f}\right] \frac{\delta h_f}{R} + \left[\frac{\sin \Phi}{\cos^2 \gamma_f} - \frac{2\bar{g}_f(R + h_f)(1 - \cos \Phi) \sin \gamma_f}{\bar{v}_f^2 \cos^3 \gamma_f}\right] \delta \gamma_f = \left[\frac{\bar{g}_f(R + h_f) \sin \Phi}{\bar{v}_f^2 \cos^2 \gamma_f} - \frac{\sin(\gamma_f + \Phi)}{\cos \gamma_f}\right] \delta \Phi - \frac{2\bar{g}_f(R + h_f)(1 - \cos \Phi)}{\bar{v}_f^3 \cos^2 \gamma_f} \delta \bar{v}_f \quad [31]$$

Considering γ_f fixed, the condition for maximum range ($\delta \Phi = 0$) becomes

$$\frac{\partial \bar{v}_f}{\partial h_f} + \frac{\bar{v}_f}{2(R + h_f)} \left[1 + \frac{\bar{v}_f^2 \cos^2 \gamma_f}{\bar{g}_f R(1 - \cos \Phi)}\right] = 0 \quad [32]$$

Alternately, assuming h_f fixed, the range is maximized when

$$\frac{\partial \bar{v}_f}{\partial \gamma_f} - \bar{v}_f \tan \gamma_f + \frac{\bar{v}_f^2 \sin \Phi}{2\bar{g}_f(R + h_f)(1 - \cos \Phi)} = 0 \quad [33]$$

Eqs. 32 and 33 are the desired end conditions for maximum ballistic range, with no restrictions on γ_f or h_f .

Orbit Altitude

Maximizing orbit altitude is of interest for ensuring the maximum orbit lifetime. Both circular and eccentric type orbits are considered in the following analysis. End conditions for the circular orbit simply state that the payload is launched horizontally at the local circular satellite velocity

$$\gamma_f = 0 \quad \bar{v}_f = \sqrt{\bar{g}_f(R + h_f)} = 0 \quad [34]$$

The optimum coasting time condition, Eq. 14, is often used in conjunction with these conditions.

For increased lifetime of an eccentric orbit, the perigee altitude must be maximized. The orbit eccentricity is specified in this application. At perigee, the radius vector has the magnitude $R + h_f = a(1 - e)$. Thus, for a specified eccentricity, maximizing the mean distance is equivalent to maximizing the perigee altitude; i.e., it is required that $\delta a = 0$ for $\delta e = 0$. Now

$$\delta a = \frac{\partial a}{\partial \bar{v}_f} \delta \bar{v}_f + \frac{\partial a}{\partial h_f} \delta h_f = 0$$

which, after substituting the derivatives, yields

$$\frac{\delta \bar{v}_f}{\delta h_f} = - \frac{\bar{g}_f}{\bar{v}_f} = \frac{\partial \bar{v}_f}{\partial h_f} + \frac{\partial \bar{v}_f}{\partial \gamma_f} \frac{\delta \gamma_f}{\delta h_f} \quad [35]$$

For specified eccentricity

$$\delta e = \frac{\partial e}{\partial \bar{v}_f} \delta \bar{v}_f + \frac{\partial e}{\partial \gamma_f} \delta \gamma_f + \frac{\partial e}{\partial h_f} \delta h_f = 0$$

which, after substituting the derivatives and writing $\delta \bar{v}_f$ in terms of δh_f , becomes

$$\frac{\delta \gamma_f}{\delta h_f} = \frac{1}{\bar{v}_f \tan \gamma_f} \left(\frac{\bar{v}_f}{R + h_f} - \frac{\bar{g}_f}{\bar{v}_f} \right) \quad [36]$$

Substituting Eq. 36 into Eq. 35

$$\left(\frac{\partial \bar{v}_f}{\partial h_f} + \frac{\bar{g}_f}{\bar{v}_f} \right) \bar{v}_f \tan \gamma_f + \left(\frac{\bar{v}_f}{R + h_f} - \frac{\bar{g}_f}{\bar{v}_f} \right) \frac{\partial \bar{v}_f}{\partial \gamma_f} = 0 \quad [37]$$

This is one of the end conditions to be satisfied. The other is

$$e_f - E = 0 \quad [38]$$

where E is the desired orbit eccentricity.

Loci of constant eccentricity form closed contours on a plot of γ_f vs. h_f . If the contour for the desired eccentricity encloses a small area, selection of appropriate initial values for the Lagrangian multipliers becomes difficult. When this situation is compounded by asking for an optimum coasting

time, the feasibility of iteration to the desired extremal is questionable. However, problems involving boost into a prescribed (mean distance, eccentricity, and inclination) orbit, with or without optimum coasting time, may be solved by employing the maximum burnout velocity option.

Hyperbolic Excess Velocity

Ascertaining the boost program that yields the maximum hyperbolic excess velocity is of interest in the direct launching of interplanetary spacecraft. From the Vis-viva integral

$$\bar{v}^2 = \bar{g}_0 R^2 \left(\frac{2}{R + h} - \frac{1}{a} \right)$$

it is apparent that the maximum velocity at infinity is achieved when the hyperbolic mean distance $-a$ is minimized; i.e., when the orbit energy is maximized. Since the orbit is conservative, this may be accomplished at burnout, where it is assumed that the path angle is specified. Setting $\delta \gamma_f = 0$ in Eq. 35, the desired end conditions become

$$\frac{\partial \bar{v}_f}{\partial h_f} + \frac{\bar{g}_f}{\bar{v}_f} = 0 \quad \gamma_f - \Gamma = 0 \quad [39]$$

A coasting period is not appropriate in this mission.

It is possible to derive analytical conditions for maximum hyperbolic excess velocity without prescribing the burnout path angle. However, as in the maximum apogee altitude mission, convergence difficulties would result.

Summary of Mission Optimizations

A summary of mission optimizations that may be achieved by applying the proposed method is given in Table 1. In practical applications, the optimum coasting time end condition, Eq. 14, would be used only with those options marked by an asterisk. Although the list of missions shown in Table 1 is fairly comprehensive, similar analyses might be employed to satisfy other specific types of mission requirements.

It was noted previously that the extremal surface represents the ultimate performance capabilities of the booster and payload specified. Obviously, the specified payload is the maximum payload with which the given booster can achieve any of the end conditions on the associated extremal surface. By applying this reasoning, the proposed method is used to ascertain the ultimate payload/mission capabilities of a given booster.

Table 1 Summary of mission optimizations

Option	Maximizes	Specify	End conditions
1*	burnout velocity	Γ, H	Eq. 12
2	hypersonic glider range	$\Gamma, W/SC_L$	Eq. 16 or 51
3	apogee altitude	Γ	Eq. 29
4	ballistic range	...	Eqs. 32, 33
5*	circular orbit altitude	$\Gamma = 0$	Eq. 34 or 52
6	elliptic orbit perigee altitude	E	Eqs. 37, 38
7	excess hyperbolic velocity	Γ	Eq. 39

Computational Techniques

Individual trajectories are computed by integration of appropriate equations of the set 4, 5, and 9, employing the given initial conditions, Eq. 11, with specified values of λ_{20}^* , λ_{30}^* and, sometimes, t_c . The desired extremal is found by iterating on these specified values until the appropriate end conditions (Table 1) are satisfied within prescribed limits. Computational aspects of this procedure, coded on an IBM 704 computer at Convair-Fort Worth, are discussed in this section.

Trajectory Integrations

All trajectory integrations are performed using the standard Runge-Kutta method. Rocket mass flow and thrust are represented by the expressions

$$\dot{m} = -\frac{1}{c}(T_R + A_{ex}P_R) \quad T = T_R + A_{ex}(P_R - P) \quad [40]$$

where T_R is a time-dependent function. The aerodynamic forces are represented by

$$C_L = (C_{L\alpha} + K|\sin \alpha|) \sin \alpha \quad C_D = C_{D0} + C_L \sin \alpha \quad [41]$$

where C_{D0} , $C_{L\alpha}$ and K are prescribed functions of Mach number. Atmospheric data are derived from subroutines representing a variety of atmosphere models.

Integration of an extremal (unconstrained) arc segment involves the simultaneous solution of Eqs. 4 and 9. In this procedure, the variable α is given by solution of the fourth of Eqs. 9 which, with the representation of Eq. 41, is a quartic having two or four real roots. For economical computer operation, it is necessary to write this equation in forms suitable for rapid iteration and selection of the extremizing root. Hence, when $\lambda_1 < 0$, values of α in the principal value range are found using

$$\alpha_j = \tan^{-1} \left\{ \frac{\lambda_2}{v\lambda_1} \left[\frac{T + (C_{L\alpha} + 2K|\sin \alpha_{j-1}|)qS}{T + qS(2C_{L\alpha} + 3K|\sin \alpha_{j-1}|) \cos \alpha_{j-1}} \right] \right\} \quad j = 1, 2, 3, 4 \quad [42]$$

where the value α_0 is taken to be either $\alpha_4 (= \alpha)$ of the previous time step or zero for the first iteration of the trajectory. When $\lambda_1 > 0$, α is found from

$$\alpha_j = \tan^{-1} \left\{ -\frac{\lambda_2}{v\lambda_1} + \frac{bS}{T} \left[(2C_{L\alpha} + 3K|\sin \alpha_{j-1}|) \sin \alpha_{j-1} - \frac{\lambda_2}{v\lambda_1} \times (C_{L\alpha} + 2K|\sin \alpha_{j-1}|) \right] \right\} \quad j = 1, 2, 3, 4 \quad [43]$$

and

$$\alpha = -\alpha_4 + \pi \frac{|\alpha_4|}{\alpha_4}$$

In Eq. 43, α_0 is either $\alpha_4 (\neq \alpha)$ of the previous time step or $-(\pi/2) \operatorname{sgn} \lambda_2$ for the first iteration. When $\lambda_1 = 0$

$$\alpha = -\frac{\pi}{2} \operatorname{sgn} \lambda_2 \quad [44]$$

Using Eqs. 42 and 43, only four iterations are required per time step to yield the accuracy level $|\Delta\alpha| \leq 0.01$ deg in the practical range of α values. When the aerodynamic forces become insignificant (arbitrarily defined by $q < 1$), their values are not computed and α is determined directly. During coasting outside of the sensible atmosphere the (ballistic) path is independent of α and its values are not computed.

In selecting the extremizing value of α , as given in Eqs. 42 through 44, it is necessary to consider the effects of excursions from the extremal. In the basic problem, consider a trajectory having fixed initial conditions and t_f , γ_f , h_f specified. The effect of local variations on the final velocity (see Eq. 7) is

given by

$$\lambda_{1f} \delta v_f = \lambda_1 \delta v + \lambda_2 \delta \gamma + \lambda_3 \delta h$$

Introducing excursions caused by variations in α , this becomes

$$\lambda_{1f}(v_f - v_f') = [\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{h} - (\lambda_1 \dot{v}' + \lambda_2 \dot{\gamma}' + \lambda_3 \dot{h}')] \Delta t$$

where v_f represents the final velocity on the extremal, v_f' is the velocity that would result by selecting the perturbed α value and Δt is a small time interval. Assuming λ_{1f} is negative, the condition for maximum final velocity is that the quantity $(\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{h})$ be a minimum; i.e.

$$\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{h} - (\lambda_1 \dot{v}' + \lambda_2 \dot{\gamma}' + \lambda_3 \dot{h}') \leq 0$$

for all possible values of α . The latter is the Weierstrass E function condition for the problem, which was discussed; e.g., in (9 and 10). Eqs. 42 through 44 are simply expressions of the Weierstrass condition in terms of the pertinent Euler equation and the aerodynamic force definition. Note that, when λ_{1f} is positive, the associated stationary value is a minimum.

For the aerodynamic force definitions of Eqs. 41, satisfying the Weierstrass condition requires that α be a continuous function of time between staging points, with two exceptions. The exceptions are the unlikely possibilities that λ_1 and λ_2 become zero simultaneously or that λ_2 becomes zero while λ_1 is positive and either the aerodynamic forces are significant or a constraint is in effect.

Arc segments constrained by one of the limits of Eq. 5 are computed like preprogrammed (nonoptimal) paths. All equations of the set 4, 5 and 9 are used, with the appropriate form of Eq. 5 solved for α , either directly or by iteration. The fourth of Eqs. 9 is used to determine λ_4 . Transition from an extremal to a constrained path occurs when one of Eqs. 5 becomes an equality. Conversely, the path again becomes an extremal when λ_4 becomes zero. Provisions are included for satisfying multiple constraints in a given problem; i.e., when two or more limits apply the computer finds and uses the most critical.

In addition to extremal and constrained segments, provisions are also made to compute constant attitude angle segments; i.e., $\gamma + \alpha = \text{constant}$, for prescribed time intervals. These are required for ground launchings since the $\dot{\gamma}$ equation

and some of the Euler-Lagrange equations are singular when the velocity is zero. Conditions at the end of constant attitude angle segments are used as initial boundary values for the variational problem spanning the remainder of the trajectory.

Basic Iteration Technique

The basic iteration procedure used to satisfy the desired end conditions was described previously (5) and only a brief summary will be given here. In this procedure the desired end conditions for an extremal containing an optimal coasting period are written as

$$\xi(v_f, \gamma_f, h_f) = 0 \quad \eta(v_f, \gamma_f, h_f) = 0 \quad [\xi(v, \gamma, h)]_{t-t^*} = 0$$

where ξ and η are the end conditions summarized in Table 1 and t^* is reserved for the optimum coasting time condition, Eq. 14. An iteration begins with the computation of four trajectories, where the first trajectory employs guessed values of λ_{20}^* , λ_{30}^* and t_c to define a reference point. The other three trajectories determine the effects on the end conditions

(in the vicinity of the reference point) of individually incrementing each of the initial conditions. Residuals produced by the values of λ_{20}^* , λ_{30}^* and t_c used are determined from the first trajectory; $R_1 = \xi$, $R_2 = \eta$, $R_3 = \zeta$.

To find the desired extremal, it is necessary that all of the residuals vanish. The first correction is based on the assumption of linear variations between the initial and final conditions, i.e.

$$|A_{sr}| \{ \Delta p_r' \} = \{ -R_s(p_r) \} \quad r = 1, 2, 3 \quad s = 1, 2, 3 \quad [45]$$

where p_1 , p_2 , p_3 are λ_{20}^* , λ_{30}^* and t_c respectively, and

$$|A_{sr}| = \begin{vmatrix} \frac{\partial \xi}{\partial \lambda_{20}^*} & \frac{\partial \xi}{\partial \lambda_{30}^*} & \frac{\partial \xi}{\partial t_c} \\ \frac{\partial \eta}{\partial \lambda_{20}^*} & \frac{\partial \eta}{\partial \lambda_{30}^*} & \frac{\partial \eta}{\partial t_c} \\ \frac{\partial \zeta}{\partial \lambda_{20}^*} & \frac{\partial \zeta}{\partial \lambda_{30}^*} & \frac{\partial \zeta}{\partial t_c} \end{vmatrix} \dots \dots \dots [46]$$

Eqs. 45 are solved to determine the next approximations to the initial values

$$p_r' = p_r + \Delta p_r' \quad [47]$$

which are used to compute the fifth trajectory.

Further corrections are determined by modifying the $|A_{sr}|$ matrix elements; i.e., the next corrections are

$$\Delta p_r'' = |A_{sr} + \Delta A_{sr}|^{-1} \{ -R_s(p_r') \} \quad [48]$$

where

$$\Delta A_{sr} = [R_s(p_r')] \Delta p_r' / \sum_{r=1}^3 (\Delta p_r')^2 \quad [49]$$

and the next approximations to the initial values are

$$p_r'' = p_r' + \Delta p_r'' \quad [50]$$

Subsequent approximations are determined in an analogous manner, with one trajectory integration per correction. The same technique, with appropriate reduction to two parameters, is used to find extremals without optimal coasting periods.

Convergence

In any iterative procedure the realization of convergence or divergence depends on the quality of the initial values used. Improvement of convergence properties implies either an increase of allowable tolerances in the initial values employed or, for a given accuracy in the initial values, a decrease in the number of iterations required to satisfy the desired end conditions. In either event, the degree of convergence has a direct bearing on the computation costs. The inherent convergence of the basic iteration technique is good. However, when this technique is adapted to finding the desired extremals, certain worth-while convergence improvements may be achieved.

For best convergence the corrections to the derivatives with respect to all three parameters, λ_{20}^* , λ_{30}^* and t_c , should proceed at similar rates. This is achieved by making all elements of a given row in the matrix of Eq. 46 have about the same magnitudes; i.e., by use of scaling factors in two columns and appropriate corrections to Eqs. 47 and 50. If no scaling factors are used, some columns never receive any significant corrections.

Trajectories with burnout at low altitudes are quite sensitive to the initial values λ_{20}^* and λ_{30}^* because of the effect of large aerodynamic forces (high dynamic pressure). This sensitivity caused some difficulty in finding extremals for hypersonic glider launchings. A significant convergence im-

provement for this mission was realized by taking the fifth root of the special end condition of Eq. 16; i.e.

$$1 - \left\{ \frac{2}{\rho} \frac{W}{SC_L} \left[\frac{\cos \Gamma}{v^2} - \frac{\cos \Gamma}{g(R+h)} - \frac{2\omega}{gv} \cos i \right] \right\}^{1/5} = 0 \quad [51]$$

Another end condition modification improved convergence of the circular satellite launching mission. In this case, the special end condition of Eq. 34 was replaced by

$$\bar{v}_f - \sqrt{\bar{g}_f(R+h_f)} + \frac{1}{4} \xi \bar{v}_f \tan \alpha_f = 0 \quad [52]$$

where $\xi = \bar{\gamma}_f = 0$ when convergence occurs. With the added term, Eq. 52 is more nearly a function of altitude alone.

The first integral of the Euler-Lagrange equations, applicable during coasting periods when the function f is independent of time, suggested another means of improving convergence.

This integral

$$\lambda_1 \dot{v} + \lambda_2 \dot{\gamma} + \lambda_3 \dot{t} = \text{constant}$$

is identical in form to the optimum coast time end condition, Eq. 14. Hence, $A_{33} \equiv 0$ in the matrix of Eq. 46 and, for most rapid convergence, this term should remain zero throughout the iteration. This is accomplished by making $\Delta A_{33} = 0$ in the matrix modification. Corrections to the other elements of the bottom row of Eq. 46 become

$$\Delta A_{3r} = \frac{[R_3(p_r')] \Delta p_r'}{(\Delta p_1')^2 + (\Delta p_2')^2} \quad r = 1, 2 \quad [53]$$

while, for $s = 1$ and 2 , Eq. 49 still applies.

During long coasting periods the Lagrangian multipliers vary significantly, amplifying the effect of poor guesses for their initial values. Usually, extreme values of α in the powered stage(s) following coasting result, often causing divergence. This situation was improved by using the constraint of Eq. 5 on $|\alpha|$ for all powered stages following coasting. The value $L_3 = 10$ deg was arbitrarily chosen.

The best way to insure that the iteration converges is to use reasonably correct values of λ_{20}^* , λ_{30}^* and t_c . To provide such values, special starting procedures are used, primarily for the first set of problems involving a given booster. A trial-and-error procedure involves specifying λ_{20}^* and t_c , then varying λ_{30}^* to find the approximate value (used in the subsequent iteration) that makes ξ become zero or reach a minimum. Since only approximate values are sought, this starting procedure uses twice the normal integration step sizes to conserve computation time. An alternate starting procedure, involving empirical relationships to define appropriate initial values, is being developed.

All of the above convergence safeguards are obviated when an input data error causes a divergent iteration. Computer time is conserved in this instance by formulating a divergence test that permits the computer to sense a divergent condition and stop. In certain other instances the iteration becomes neutrally convergent; i.e., the individual trajectory solutions oscillate about the desired end conditions. A special test allows the computer to sense this condition and to restart the iteration, beginning with the best trajectory of those computed up to the point where the test is made.

Numerical Results

The method described has been employed to compute numerous extremals for both air and ground launched boost vehicles. Two ground launched extremals, exhibiting some interesting properties, have been selected for illustrative purposes.

Table 2 Boost rocket characteristics
Thrust and Weight Data

	Stage 1		Stage 2		Stage 3	
	A	B	A	B	A	B
Rated thrust, lb $\times 10^{-3}$	840 ^a	336 ^a	140 ^b	70 ^b	15.12 ^b	7.56 ^b
Burning time, sec	60	150	60	120	60	120
Nozzle exit area, ft ²	25	10	4	2
Specific im- pulse, sec	250 ^a		280 ^b		280 ^b	
Stage weight, lb $\times 10^{-3}$	219.1		32.6		3.6	
Propellant weight, lb. $\times 10^{-3}$	201.6		30.0		3.24	

Stage 1 Aerodynamic Data^c

Mach			Mach		
No.	C_{D0}	C_{La}	No.	C_{D0}	C_{La}
0	0.425	11.6	3.0	0.791	9.4
0.5	0.425	11.8	6.0	0.609	8.1
0.9	0.425	13.3	10.0	0.560	8.0
1.2	1.185	14.5	15.0	0.560	7.3

^a Sea level.

^b Vacuum.

^c $K = 0$, $S = 70$ ft², ARDC 1956 atmosphere.

Boost Rockets

Two rocket vehicles, designated as rockets A and B were used in the computed examples. Pertinent characteristics of these vehicles are shown in Table 2. The total impulse of the two rockets is identical. Aerodynamic forces during second stage burning were computed using the coefficients for the first stage, with $S = 30$ ft².

Extremal Comparisons

Extremals for a maximum altitude circular, polar orbit were computed for the two rockets. In both cases the orbit altitude attained was 227 n mi. Time histories of the two extremals are shown in Fig. 2, with the associated payloads and percentages of ideal velocity realized. Altitude effects on thrust, along the respective extremal, were included in the calculation of the ideal velocity and the percentages realized include losses during coasting. Optimum coasting times for rockets A and B were 416 and 309 sec, respectively. Initial pitch times of 10 and 25 sec, yielding nearly the same pitch altitudes, were selected arbitrarily.

The optimum pitch programs during the first two stages provide an interesting contrast. Rocket A, having a relatively short burning time required a continuous pitching program (negative α) to best accomplish the mission. Rocket B achieved optimum performance by pitching rapidly early in the flight, then continuing to second stage burnout at positive α values. In this manner, the gravity losses were effectively reduced. In regions of high dynamic pressure the magnitude of α decreased on both extremals, to reduce the aerodynamic drag losses. At the times of maximum dynamic pressure, the angles of attack of rockets A and B were 5.8 and 1.4 deg, respectively. Last stage extremal arcs for both rockets were similar to the approximate vacuum program derived by Fried (11).

Extremal Simulation

Two comments noted previously—that discontinuities in α have a minor effect on the entire trajectory and that extremals had minimal sensitivity to guidance and control errors—implied that extremals could readily be simulated by practical guidance and control techniques. In the interest of verifying this hypothesis and ascertaining the order of magnitude of performance losses involved, the extremal of rocket A was simulated as indicated in Fig. 3. The simulation consisted of four constant attitude angle rate segments in the powered phases, with time spans and magnitudes as noted on the figure, a program requiring horizon sensors for guidance. In this example, the burnout altitude and path angle were fixed and the simulated trajectory had a burnout velocity of 56 fps less than the extremal. Larger velocity discrepancies would result for poorer approximations to the extremal arc.

Conclusions

The flexibility of the indirect methods of the calculus of variations in performing mission optimizations for rocket boost vehicles has been demonstrated. This method has proved feasible for vehicle evaluations during design studies, primarily because of the ability to achieve adequate vehicle and mission simulation without excessive computer costs.

A further comment regarding optimizations for various classes of vehicles and missions is appropriate. The employment of variational techniques to achieve such optimizations provides the most systematic basis for making logical design decisions. On such a basis, the design requirements of a vehicle and its subsystems may be established with increased confidence.

Acknowledgment

The authors are indebted to L. G. Walker of the Convair-Fort Worth Computation Laboratory for the computer programming of the problem.

Nomenclature

A_{ex}	= nozzle exit area
C_D	= drag coefficient, D/qS
C_L	= lift coefficient, L/qS
c	= effective exhaust velocity
D	= drag force
g	= acceleration of gravity
h	= altitude
i	= inclination of trajectory plane
L	= lift force
L_m	= limit value of control function
l_m	= control functions for constrained paths
P	= ambient pressure
P_R	= ambient pressure at rated altitude
q	= dynamic pressure, $(1/2)\rho v^2$
R	= planet radius
S	= reference area
T	= thrust
T_{eq}	= equilibrium wall temperature
T_R	= rated thrust
t	= time
v	= velocity
W	= instantaneous vehicle weight
α	= angle of attack and angle between thrust and velocity vectors
β	= angle between constrained acceleration vector and normal to vehicle longitudinal axis, measured positive aft of reference
γ	= flight path angle, referenced to local horizontal
ρ	= atmospheric density
ω	= rotation rate of planet

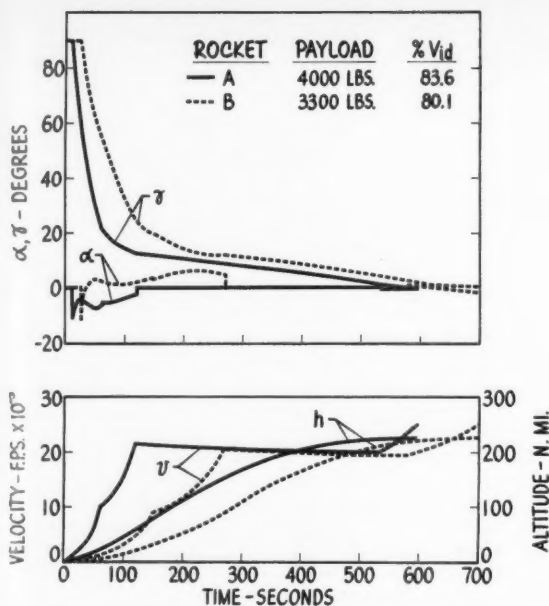


Fig. 2 Comparison of satellite launching extremals

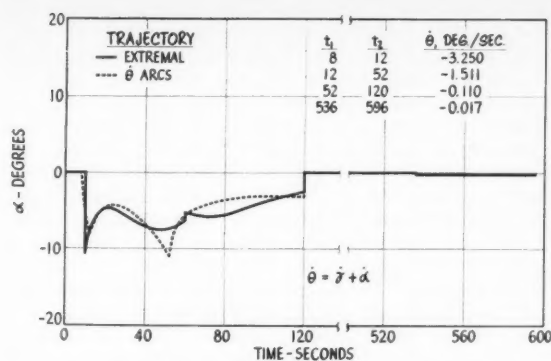


Fig. 3 Simulation of rocket A extremal

References

- 1 Miele, A., "General Variational Theory of the Flight Paths of Rocket-Powered Aircraft, Missiles and Satellite Carriers," *Astro. Acta*, vol. 4, 1958, pp. 264-288.
- 2 Miele, A., "A Survey of the Problems of Optimizing Flight Paths of Aircraft and Missiles," ARS preprint 1219-60.
- 3 Breakwell, J. V., "The Optimization of Trajectories," *J. Soc. Indust. Appl. Math.*, vol. 7, no. 2, June 1959, pp. 215-247.
- 4 Kelley, H. J., "Gradient Theory of Optimal Flight Paths," ARS preprint 1230-60.
- 5 Kulakowski, L. J. and Stancil, R. T., "Rocket Boost Trajectories for Maximum Burn-Out Velocity," *ARS JOURNAL*, vol. 30, no. 7, July 1960, pp. 612-618.
- 6 Bliss, G. A., *Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, First Ed., 1946, Chap. 7.
- 7 Cieala, C. and Miele, A., "Brachistochronic Maneuvers of a Variable Mass Aircraft in a Vertical Plane," *J. Aero. Sci.*, vol. 22, no. 8, Aug. 1955, pp. 577-578.
- 8 Wheelon, A. D., "Free Flight of a Ballistic Missile," *ARS JOURNAL*, vol. 29, no. 12, Dec. 1959, pp. 915-926.
- 9 Leitmann, G., "On a Class of Variational Problems in Rocket Flight," *J. Aero/Space Sci.*, vol. 26, no. 9, 1959.
- 10 Leitmann, G., "Minimum Transfer Time for a Power-Limited Rocket," XI Int. Astronautical Congress, Stockholm, 1960.
- 11 Fried, B. D., "On the Powered Flight Trajectory of an Earth Satellite," *JET PROPULSION*, vol. 27, no. 6, June 1957, pp. 641-643.

Optimum Maneuvers for Launching Satellites Into Circular Orbits of Arbitrary Radius and Inclination

J. P. CARSTENS¹ and
T. N. EDELBAUM²

United Aircraft Corporation
Research Laboratories
East Hartford, Conn.

A method of launching an Earth satellite into a circular orbit that does not pass over its launch point on a stationary Earth is described. The path followed is a direct ascent using only two thrust impulses so that the satellite vehicle follows a single ballistic arc from the first impulse at launch to the second impulse at injection into the target orbit. The minimum propulsion requirements for this type of trajectory are found as a function of the target orbit to launch-point radius ratio and of the inclination angle between the radius vector to the launch point and its orthogonal projection on the plane of the target orbit. This maneuver is compared with more conventional satellite launching methods which generally involve an initial parking orbit.

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ Analytical Research Engineer. Member ARS.

² Senior Research Engineer. Member ARS.

ESTABLISHED methods of launching a satellite into an orbit of arbitrary inclination to some reference plane usually consist of first entering a low altitude circular parking orbit which passes directly over the launch point and then using a minimum fuel transfer from this parking orbit to the desired final orbit, e.g., (1 and 2).³ The assumption is made either implicitly (1) or explicitly (2) that the optimum trajectory for achieving a high altitude inclined orbit employs a parking orbit. One way of testing this assumption, and thereby establishing its validity for future mission analyses, is to compare it with the simplest maneuver that does not utilize a parking orbit, namely a direct ascent using two thrust impulses. The simplicity of the direct ascent maneuver allows a complete optimization of this type of trajectory and provides a basic reference against which more complex maneuvers, with or without parking orbits, can be evaluated. Although the study is primarily concerned with idealized propulsion requirements, there is some discussion of the relative operational advantages of maneuvers utilizing parking orbits.

The study reported herein is an outgrowth of one aspect of a report concerning interceptor satellites (3). In (3), Edelbaum discusses the significant savings in total characteristic velocity that may be realized for certain conditions by the use of optimum direct ascent trajectories. The purpose of this paper is to present in more detail the characteristics of optimum two-impulse trajectories for the general case of ascent to any inclined circular orbit and to compare these trajectories with indirect ascents which use a parking orbit.

Assumptions

The analysis of this paper, which is given in detail in Appendixes I and II, is carried out for a nonrotating spherical Earth. Atmospheric forces are neglected, and all applications of thrust are assumed to be impulsive. These results may be applied to more realistic cases by assuming that the nominal launch point lies slightly above the atmosphere of the planet, and by adding gravity and drag losses to the required characteristic velocity to account for the effects of the atmosphere and of the finite burning time. As these losses are relatively insensitive to the trajectory for reasonably high thrust to weight ratios, an analysis such as the present one can provide a good approximation to more complete analyses carried out for specific vehicles.

The geometry of a typical two-impulse transfer is shown in Fig. 1. The first impulse is applied at the launch point L and the second at the intersection with the target orbit P . The target orbit is defined by its radius r_2 and by its inclination i to the launch point. The inclination is the angle between the radius vector to the launch point and its orthogonal projection on the plane of the target orbit. For the special case of an equatorial target orbit, the inclination would correspond to the latitude of the launch point. For a specific target orbit, only two independent variables are needed to specify the transfer trajectory; the two used herein are the range angle ϕ from launch to interception of the target orbit, and the launch angle θ_1 with respect to the local horizontal.

Optimum trajectories are considered to be those requiring a minimum total characteristic velocity to achieve a given orbit. Minimum values of total characteristic velocity (Eq. 5 of Appendix I), and the minimizing values of the two independent variables were found directly on a high speed digital computer.

The minimum characteristic velocity criterion will maximize payload for rockets which are restartable. Restartable rockets allow the thrust impulses, i.e., burning periods, to be distributed independently of the staging arrangement, and

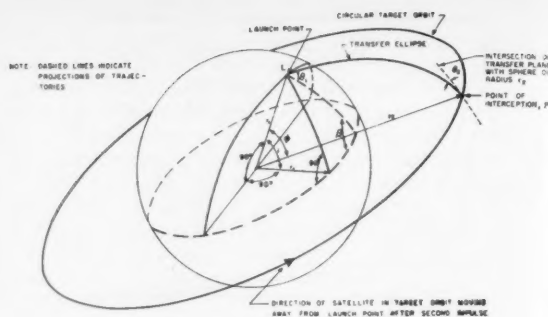


Fig. 1 Two-impulse trajectory geometry

the staging optimization may thereby remain distinct from the trajectory optimization. Nonrestartable rockets require an integral number of stages to be burned for each impulse, and thus necessitate consideration of the interaction of the staging with the impulse distribution.

Results of Trajectory Optimization

Minimum characteristic velocity requirements, nondimensionalized by circular velocity at r_1 and therefore applicable to any size planet, are presented in Fig. 2 for a range of values of i as a function of r_2/r_1 . These minimum values were determined by numerical methods on an IBM 704 computer. As is known for satellite launching when $i = 0$ deg, e.g., (4), the required characteristic velocity reaches a maximum for r_2/r_1 of about 15.6, thereafter decreasing toward an asymptote of $\sqrt{2}$ times circular velocity as r_1 goes to infinity. The target orbit radius corresponding to the maximum velocity requirement decreases as the inclination of the target orbit plane increases until, for i greater than about 45 deg, the most difficult orbit to reach has the lowest possible altitude.

The corresponding optimum trajectory variables are shown in Fig. 3. A value of ϕ of 90 deg represents the case where the angle between the plane of the transfer orbit and the plane of the target orbit β reaches its minimum value which is equal to i . This condition minimizes one component of the second velocity impulse and might be considered as an approximation to the optimum value of ϕ . Fig. 3 shows, however, that this effect does not predominate, and in general the optimum ϕ is not equal to 90 deg. As would be expected, the use of trajectories with $\phi = 90$ deg does provide a good approximation to the true optimum when i is large, but does not provide a good approximation when i is small. One interesting feature of the curves of Fig. 3 is that for $i = 0$ the optimum ϕ always equals 180 deg (Hohmann ellipse) except for the case $r_2/r_1 = 1$. For this special case, ϕ is meaningless since within the assumptions of the analysis θ_{1OPT} is zero, and the total characteristic velocity is equal to local circular velocity. However, by assuming θ and i to be small angles, it is shown in Appendix II that ϕ_{OPT} approaches 70.529 deg as i goes to zero, as indicated in Fig. 3.

Another trajectory condition that seems intuitively advantageous is the location of the apogee of the transfer arc on the target orbit, which would obviate any radial velocity correction at the second impulse. This approximation has been considered in (5). Comparison of the results of (5) with the optimum results of the present paper shows that this approximation is poor at low radius ratios. There does not seem to be any simple approximation to the optimum two-impulse transfer which is valid over the entire range of values of r_2/r_1 and i .

³ Numbers in parentheses indicate References at end of paper.

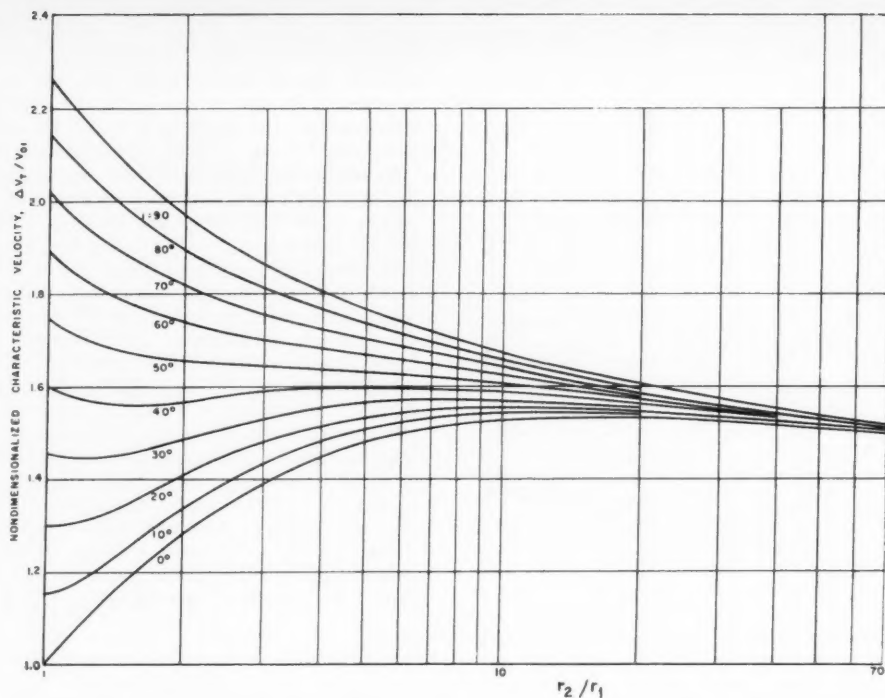


Fig. 2 Minimum characteristic velocity requirements of two-impulse trajectory

Comparison With Other Maneuvers

The primary purpose of the present paper is to compare direct ascent trajectories with indirect ascents which use a parking orbit. The indirect ascent considered, entry into a circular parking orbit at essentially the original launch radius followed by the inclined Hohmann transfer of (6), is used in (1) and (2). This indirect ascent requires three impulses. The two maneuvers are compared in nondimensional form in Fig. 4. It can be seen that the optimum two-impulse transfer is superior to this particular three-impulse transfer for all inclinations when the radius ratio is small and is inferior for almost all inclinations when the radius ratio is large. The percentage differences in characteristic velocity range only up to about 6%, but the absolute differences are more significant in terms of vehicle payload capability. Fig. 5 presents these differences between the characteristic velocity requirements of the two maneuvers, taking the nominal initial radius at 200 n mi above the surface of the Earth. Negative values of the difference represent cases where the two-impulse transfer requires more characteristic velocity, and positive values indicate the reverse. It can be seen that the optimum two-impulse transfer never increases the characteristic velocity requirements by more than about 650 fps whereas it can decrease these requirements by as much as 3800 fps. The significance of these results is demonstrated by an example in (3), where a 3800 fps decrease in required characteristic velocity increases payload approximately 50%.

The reason for the significant advantage of the direct ascent trajectory at large inclinations can be demonstrated by considering the somewhat academic case of launching from the North Pole into an equatorial orbit. In this case the transfer trajectory can have no component in the direction of motion of the final orbit, so that the characteristic velocity is minimized by minimizing the velocity at both ends of the transfer arc. Since a minimum energy trajectory will have this

property, the optimum two-impulse transfer when the inclination is 90 deg is a minimum energy transfer from the launch point to a point on the final orbit. For a low altitude orbit, the optimum direct ascent would have a velocity of only about 23,000 fps (typical of an ICBM) at each end of the transfer arc, whereas a circular parking orbit would have a velocity of about 25,000 fps at each end of the transfer arc. This is perhaps the clearest example of why parking orbits are not always desirable.

There may be operational advantages to parking orbits, as compared with direct ascents, but a preliminary examination of this problem has not revealed any conspicuous advantages. In most cases the launch azimuths for the two maneuvers are not very different so that range safety, tracking acquisition, and Earth spin effects should not be too different. If necessary, the indirect ascent can use exactly the same launch azimuth as the parking orbit ascent with never more than a 5% penalty in characteristic velocity. For large inclinations, where direct ascents are particularly desirable, this penalty becomes negligibly small. Timing considerations (for changing epoch angle) for low altitude parking orbits fall outside the scope of the present paper, but (1) and (2) indicate that even here, the low altitude parking orbit may not be especially desirable.

Perhaps the most serious disadvantage to the direct ascent scheme is the larger launch angle required, which will cause larger gravity losses compared to entering a parking orbit, particularly for vehicles with low thrust-to-weight ratios.

Twenty Four Hour Orbit

One of the more important Earth satellite missions is launching a 24-hr equatorial satellite from a nonequatorial launch site. For Earth this corresponds to an r_2/r_1 of about 6.25. Reference to Fig. 4 shows that for this radius ratio there is very little difference between the optimum two-

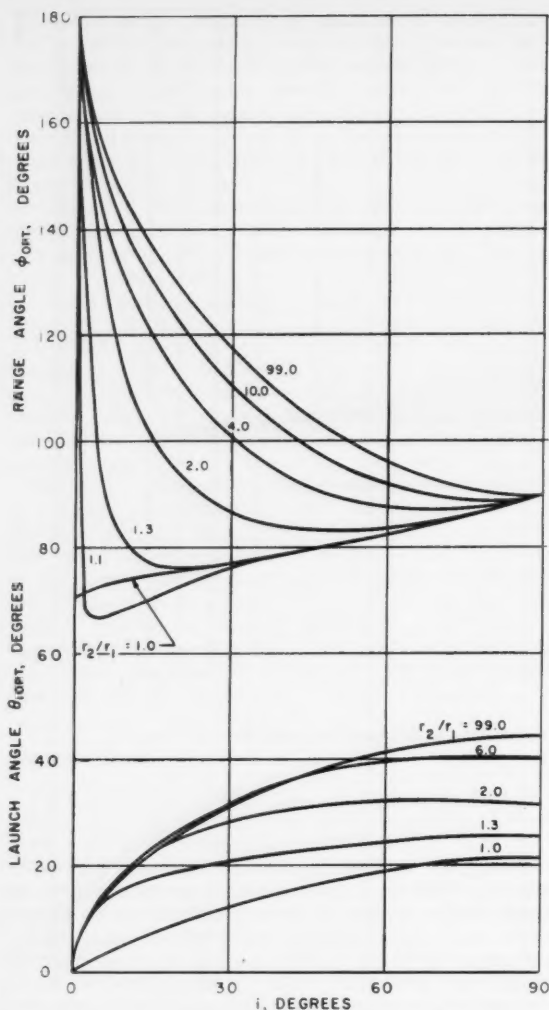


Fig. 3 Range and launch angles for optimum two-impulse trajectories

impulse trajectory and the three-impulse indirect ascent. At an inclination of 28 deg, corresponding to launching an equatorial satellite from Cape Canaveral, the difference in characteristic velocity is only about 500 fps. Although for this particular case the two-impulse maneuver requires slightly more characteristic velocity than the three-impulse maneuver, its simplicity may make it desirable. Any desired longitude could be achieved by using the technique referred to in (2) as "high-altitude looping," which requires one extra impulse. Since "high-altitude looping" appears to be desirable even when a parking orbit is used, the direct ascent would still require one less impulse than would the parking orbit trajectory.

Optimum Three-Impulse Trajectories

Since the two-impulse transfer is often superior to the three-impulse transfers considered in the literature, the truly optimum three-impulse transfer should require even less fuel than either trajectory considered here. In particular, the use of this extra impulse would probably produce significant decreases in characteristic velocity for cases where the change in inclination is large. This result is indicated by the im-

provement in the three-impulse transfer when a four-impulse bi-elliptic transfer is utilized (1). Another indication of the advantage of the extra impulse is given in (7) where a transfer between inclined circular orbits with the same radius was considered. At the present time, the optimum maneuver with more than two impulses, as well as the optimum number of impulses for any particular case, remains unsolved.

However, the available evidence is sufficient to determine that the optimum three-impulse maneuver will generally not involve a low altitude parking orbit. This report assumes that only one impulse is used to enter a parking orbit. Since the inclined Hohmann transfer is apparently the optimum two-impulse transfer from one circular orbit to another, the three-impulse transfers considered herein are apparently the optimum three-impulse transfers using a low altitude parking orbit. However, the true optimum three-impulse transfer would never require more characteristic velocity than the optimum two-impulse transfer. Because two of the impulses may be given successively, the three-impulse maneuver may always degenerate into a two-impulse maneuver. Thus, when we speak of a three-impulse maneuver, we mean a "not more than three-impulse" maneuver. Since the best three-impulse transfer using a parking orbit is often inferior to the two-impulse transfer, it can be concluded that the true optimum three-impulse transfer would generally not involve entry into a parking orbit.

Conclusions

- 1 Minimum fuel maneuvers for entering inclined circular orbits generally do not involve low altitude parking orbits.
- 2 The use of optimum two-impulse maneuvers for direct

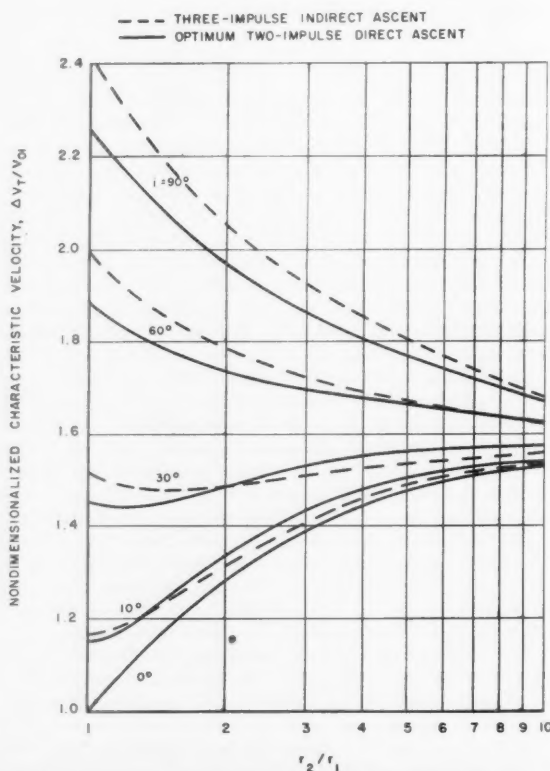


Fig. 4 Comparison of direct and indirect ascent trajectories—nondimensional characteristic velocity requirements

ascent to inclined circular orbits can result in significant reductions in required characteristic velocity relative to conventional indirect ascent maneuvers which involve a low altitude parking orbit.

Appendix I

Two-Impulse Characteristic Velocity Requirements

The first velocity impulse is given simply by

$$\frac{\Delta V_1}{V_{01}} = \sqrt{2 \frac{\mu}{r_1} - \frac{\mu}{a}} = \sqrt{2 - \frac{r_1}{a}} \quad [1]$$

where a is the semi-major axis of the transfer ellipse from L to P (Fig. 1).

At P , the projection of the transfer trajectory on the sphere r_2 makes an angle β with the plane of the target orbit, and a line tangent to the transfer trajectory at P forms an angle θ_2 with the local horizontal at P . The second velocity impulse is then

$$\frac{\Delta V_2}{V_{01}} = \sqrt{\frac{r_1}{r_2} - 2 \sqrt{\frac{r_1}{r_2} \frac{V_2}{V_{01}}} \cos \theta_2 \cos \beta + \left[\frac{V_2}{V_{01}} \right]^2} \quad [2]$$

But

$$\cos \beta = \frac{\sqrt{\sin^2 \phi - \sin^2 i}}{\sin \phi} \quad [3]$$

and

$$\left[\frac{V_2}{V_{01}} \right]^2 = 2 \frac{r_1}{r_2} - \frac{r_1}{a} \quad [4]$$

Therefore

$$\frac{\Delta V_T}{V_{01}} = \frac{\Delta V_1 + \Delta V_2}{V_{01}} = \sqrt{2 - \frac{r_1}{a}} + \sqrt{\frac{r_1}{r_2} - 2 \sqrt{\frac{r_1}{r_2} \left(2 \frac{r_1}{r_2} - \frac{r_1}{a} \right)} \cos \theta_2 \frac{\sqrt{\sin^2 \phi - \sin^2 i}}{\sin \phi} + 2 \frac{r_1}{r_2} - \frac{r_1}{a}} \quad [5]$$

Minimum values of $\Delta V_T/V_{01}$ were found as a function of a/r_1 and ϕ for given sets of r_1/r_2 and i . The optimization was done numerically on an IBM 704 computer. Note that for given r_1/r_2 and i

$$i \leq \phi \leq 180 \text{ deg} - i \quad \text{and} \quad \frac{a}{r_1} \geq \left[\frac{a}{r_1} \right]_{\text{MIN}} =$$

$$\frac{1 + \frac{r_2}{r_1} + \sqrt{1 + \left(\frac{r_2}{r_1} \right)^2 - 2 \frac{r_2}{r_1} \cos \phi}}{4} \quad [6]$$

For a given selection of a/r_1 and ϕ (8)

$$\sec^2 \theta_2 = \frac{r_2}{l} \left(2 - \frac{r_2}{a} \right) \quad [7]$$

where

$$\frac{l}{r_1} = \left\{ \frac{4 \frac{a}{r_1} \left\{ \left[2 \left(\frac{a}{r_1} \right)_{\text{MIN}} - 1 \right] \left[2 \left(\frac{a}{r_1} \right)_{\text{MIN}} - \frac{r_2}{r_1} \right] \right\}}{1 + \left(\frac{r_2}{r_1} \right)^2 - 2 \frac{r_2}{r_1} \cos \phi} \right\} \times \sin^2 \left[\frac{\alpha \pm \beta}{2} \right] \quad [8]$$

$$r_1 = 3640 \text{ N MI}$$

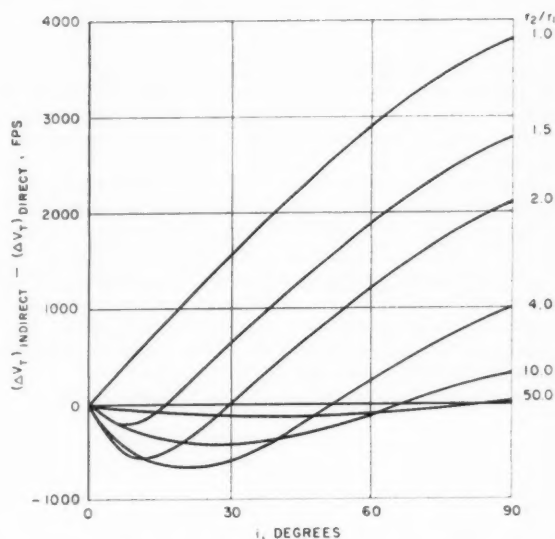


Fig. 5 Difference in required characteristic velocity between direct and indirect ascent trajectories

where

$$\sin \frac{\alpha}{2} = \sqrt{\frac{a_{\text{MIN}}}{a}} \quad [9]$$

and

$$\sin \frac{\beta}{2} = \sqrt{\frac{\frac{1}{2} + \frac{1}{2} \frac{r_2}{r_1} - \frac{1}{2} \frac{c}{r_1}}{2 \frac{a}{r_1}}} \quad [10]$$

Appendix II

Special Case of $r_1 = r_2$

For the case of $r_1 = r_2$, considering the launch angle θ_1 and the range angle ϕ as the independent variables, it is easily shown that

$$\frac{\Delta V_1}{V_{01}} = \sqrt{2 - \frac{r_1}{a}} = \frac{1}{\cos \theta_1 \sqrt{1 + \frac{\tan \theta_1}{\tan \frac{\phi}{2}}}} \quad [11]$$

and since $\Delta V_1 = V_2$, then

$$\frac{\Delta V_1}{V_{01}} + \frac{\Delta V_2}{V_{01}} = \frac{1}{\cos \theta_1 \sqrt{1 + \frac{\tan \theta_1}{\tan \frac{\phi}{2}}}} + \sqrt{1 - 2 \cos \beta \frac{\cos \theta_2}{\cos \theta_1 \sqrt{1 + \frac{\tan \theta_1}{\tan \frac{\phi}{2}}}}} + \frac{1}{\cos^2 \theta_1 \left(1 + \frac{\tan \theta_1}{\tan \frac{\phi}{2}} \right)}$$

but, for $r_2 = r_1$, $\theta_1 = -\theta_2$, so $\cos \theta_2 = \cos \theta_1 = \cos \theta$ ($\theta < 90$ deg)

and so

$$\frac{\Delta V_T}{V_{01}} = \frac{1}{\cos \theta \sqrt{1 + \frac{\tan \theta}{\tan \frac{\phi}{2}}}} + \sqrt{1 - \frac{2\sqrt{1 - \frac{\sin^2 i}{\sin^2 \phi}}}{\sqrt{1 + \frac{\tan \theta}{\tan \frac{\phi}{2}}}}} + \frac{1}{\cos^2 \theta \left(1 + \frac{\tan \theta}{\tan \frac{\phi}{2}}\right)} \quad [12]$$

From the numerical results it is apparent that as i approaches 0, ϕ_{OPT} does not approach 180 deg. Since this is in direct variance with the results of the general case, $r_2/r_1 > 1$ (Fig. 3), it is of interest to define ϕ_{OPT} for this special case of small i and $r_2 = r_1$.

Defining this point graphically by repeatedly solving the above equation would be tedious due to the generally slight variation of $\Delta V_T/V_{01}$ with ϕ . Sample calculations show that $\Delta V_T/V_{01}$ does rise sharply, however, for ϕ near i or 180 deg $- i$, since approaching these values of ϕ requires intercepting the target orbit at angles which approach 90 deg for any value of i . Neglecting values of ϕ close to i (where θ_{OPT} approaches 45 deg), it is valid to assume that θ_{OPT} becomes a small angle as i becomes small.

Dropping θ and i terms smaller than first order gives

$$\cos \theta = 1, \tan \theta = \theta, \text{ and } \sin i = i$$

$$\frac{\Delta V_T}{V_{01}} = \frac{1}{\sqrt{1 + \frac{\theta}{\tan \frac{\phi}{2}}}} + \sqrt{1 - \frac{2\sqrt{1 - \frac{i^2}{\sin^2 \phi}}}{\sqrt{1 + \frac{\theta}{\tan \frac{\phi}{2}}}}} + \frac{1}{(1 - \theta^2) \left(1 + \frac{\theta}{\tan \frac{\phi}{2}}\right)} \quad [13]$$

where second order terms must be kept under the radical sign. Then

$$\frac{\Delta V_T}{V_{01}} = \left[1 - \frac{\theta}{2 \tan \frac{\phi}{2}}\right] + \sqrt{1 - 2 \left(1 - \frac{i^2}{2 \sin^2 \phi}\right) \left(1 - \frac{\theta}{2 \tan \frac{\phi}{2}} + \frac{3}{8} \frac{\theta^2}{\tan^2 \frac{\phi}{2}}\right)} + (1 + \theta^2) \left(1 - \frac{\theta}{\tan \frac{\phi}{2}} + \frac{\theta^2}{\tan^2 \frac{\phi}{2}}\right)$$

which simplifies to

$$\frac{\Delta V_T}{V_{01}} = 1 - \frac{\theta}{2 \tan \frac{\phi}{2}} + \sqrt{\frac{i^2}{\sin^2 \phi} + \frac{\theta^2}{4 \tan^2 \frac{\phi}{2}}} + \theta^2 \quad [14]$$

To solve for ϕ_{OPT} will require expressions containing both ϕ_{OPT} and θ_{OPT} . Setting the respective derivatives equal to zero yields

$$\frac{\partial \frac{\Delta V_T}{V_{01}}}{\partial \phi} = \frac{\theta}{4 \sin^2 \frac{\phi_{OPT}}{2}} - \frac{\left[\frac{2i^2 \cos \phi_{OPT}}{\sin^3 \phi_{OPT}} + \frac{\theta^2 \cos \frac{\phi_{OPT}}{2}}{4 \sin^3 \frac{\phi_{OPT}}{2}} \right]}{2 \sqrt{\frac{i^2}{\sin^2 \phi_{OPT}} + \frac{\theta^2}{4 \tan^2 \frac{\phi_{OPT}}{2}}} + \theta^2} = 0 \quad [15]$$

and

$$\frac{\partial \frac{\Delta V_T}{V_{01}}}{\partial \theta} = -\frac{1}{2 \tan \frac{\phi}{2}} + \frac{\frac{\theta_{OPT}}{2 \tan^2 \frac{\phi}{2}} + 2\theta_{OPT}}{2 \sqrt{\frac{i^2}{\sin^2 \phi} + \frac{\theta_{OPT}^2}{4 \tan^2 \frac{\phi}{2}}} + \theta_{OPT}^2} = 0 \quad [16]$$

Solving the latter equation for θ_{OPT} results in

$$\theta_{OPT}^2 = \frac{i^2}{\sin^2 \phi \left(4 \tan^2 \frac{\phi}{2} + 1\right)} \quad [17]$$

Now rewriting Eq. 15 for ϕ_{OPT}

$$\theta \left\{ \frac{1}{4 \sin^2 \frac{\phi_{OPT}}{2}} - \frac{\left[\frac{2i^2 \cos \phi_{OPT}}{\sin^3 \phi_{OPT}} + \frac{\cos \frac{\phi_{OPT}}{2}}{4 \sin^3 \frac{\phi_{OPT}}{2}} \right]}{2 \sqrt{\frac{i^2}{\sin^2 \phi_{OPT}} + \frac{1}{4 \tan^2 \frac{\phi_{OPT}}{2}}} + 1} \right\} = 0 \quad [18]$$

or, upon substituting Eq. 17 for i/θ_{OPT}

$$\theta_{OPT} \left\{ \frac{1}{4 \sin^2 \frac{\phi_{OPT}}{2}} - \frac{\left(\frac{4 \tan^2 \frac{\phi_{OPT}}{2} + 1 \right)}{\tan \phi_{OPT}} + \frac{1}{8 \sin^2 \frac{\phi_{OPT}}{2} \tan \frac{\phi_{OPT}}{2}} \right\} = 0$$

$$\frac{1}{2 \tan \frac{\phi_{OPT}}{2}} \sqrt{16 \tan^4 \frac{\phi_{OPT}}{2} + 8 \tan^2 \frac{\phi_{OPT}}{2} + 1} = 0$$

which reduces to

$$\theta_{\text{OPT}} \left\{ \frac{1}{\sin^2 \frac{\phi_{\text{OPT}}}{2}} \left[\frac{\tan^2 \frac{\phi_{\text{OPT}}}{2}}{1 + 4 \tan^2 \frac{\phi_{\text{OPT}}}{2}} \right] - \frac{2 \tan \frac{\phi_{\text{OPT}}}{2}}{\tan \phi_{\text{OPT}}} \right\} = 0 \quad [19]$$

Neglecting the solution $\theta_{\text{OPT}} = 0$, which occurs only when $i = 0$ and ϕ_{OPT} is indeterminate, the function within the brackets reduces to

$$3 \cos^2 \phi_{\text{OPT}} - 4 \cos \phi_{\text{OPT}} + 1 = 0 \quad [20]$$

which has roots at $\cos \phi_{\text{OPT}} = 1$ and $\cos \phi_{\text{OPT}} = \frac{1}{3}$. Ignoring the first root ($\phi_{\text{OPT}} = 0$, which is possible only when $i = 0$), $\cos \phi_{\text{OPT}} = \frac{1}{3}$, and ϕ_{OPT} for this case is 70.529 deg.

These optimum values result in the following approximate expression for the required characteristic velocity, valid for small i

$$\frac{\Delta V_T}{V_{01}} = 1 + \frac{\sqrt{3}}{2} i = 1 + 0.8660i \quad [21]$$

The case of $\theta = 0$ and $\phi = 90$ deg, corresponding to entry into the best circular parking orbit, would result in a characteristic velocity of

$$\frac{\Delta V_T}{V_{01}} = 1 + i \quad [22]$$

Finally, if the orbit was entered at the minimum or maximum allowable values of ϕ (i , $180 - i$) the minimum required characteristic velocities for small i would be respectively

$$\frac{\Delta V_T}{V_{01}} = 1 + \sqrt{i} \quad [23]$$

$$\frac{\Delta V_T}{V_{01}} = 1 + \sqrt{2} = 2.414 \quad [24]$$

Nomenclature

a = semi-major axis of transfer ellipse

c = shortest distance from L to P
 i = angle between radius vector to launch point, and its orthogonal projection on the target orbit plane
 L = launch point, position of first impulse
 l = semi-latus rectum of transfer ellipse
 P = point of interception with target orbit
 r_1 = radius of launch sphere
 r_2 = radius of target orbit
 V_{01} = circular velocity at r_1
 V_2 = velocity in the transfer ellipse at P
 ΔV_1 = first velocity impulse
 ΔV_2 = second velocity impulse
 $\Delta V_T = \Delta V_1 + \Delta V_2$
 α = see Eq. 9, Appendix I
 β = angle between transfer ellipse and target orbit projected on sphere r_1 . Also as defined in Appendix, Eq. 10
 θ_1 = angle between local horizontal and tangent to flight path at launch
 θ_2 = angle between local horizontal and tangent to flight path at intercept, see Fig. 1
 ϕ = range angle: angle between launch point, center of Earth, and point of interception
 μ = gravitational constant for a body

Subscripts

MIN = minimum value

OPT = independent variable value for which required characteristic velocity is a minimum

References

1. Fiul, A. and Braham, H., "Optimization of Vehicles and Trajectories for the Twenty-Four Hour Equatorial Satellite Mission," ARS preprint 1120-60, May 1960.
2. Hoelker, R. and Silber, R., "Analysis of Injection Schemes for Obtaining a Twenty-Four Hour Orbit," IAS paper no. 60-75, June 1960.
3. Edelbaum, T. N., "Preliminary Comparison of Air and Ground Launching of Satellite Rendezvous Vehicles," IAS paper no. 61-10, Jan. 1961.
4. Lawden, D. F., "Entry into Circular Orbits," *J. Brit. Interplanet. Soc.* vol. 10, no. 5, 1951.
5. Wolfe, J. F., and DeBra, D., "Two Maneuver Ascents to Circular Orbits," *J. Astron. Sci.*, vol. VII, no. 2, Summer 1960.
6. Horner, J., and Silber, R., "Impulse Minimization for Hohmann Transfer Between Inclined Circular Orbits of Different Radii," Army Ballistic Missile Agency, Rep. no. DA-TR-70-59, Dec. 1959.
7. Edelbaum, T. N., "Propulsion Requirements for Controllable Satellites," ARS preprint 1228-60, May 1960.
8. Battin, R. H., "The Determination of Round-Trip Planetary Reconnaissance Trajectories," *J. Aero/Space Sci.*, vol. 26, no. 9, Sept. 1959.

Self-Contained Navigational System for Determination of Orbital Elements of a Satellite

K. N. SATYENDRA¹
and R. E. BRADFORD²

Nortronics, A Div. of
Northrop Corp.
Hawthorne, Calif.

With the increase in satellite traffic, a requirement exists for a self-contained or on-board navigational system which can determine the six orbital elements of a satellite. This paper develops the detailed mathematical techniques for the updating of the orbital elements of a satellite. The approach used is essentially the method of differential corrections applied to on-board measurements made of the local vertical by a horizon tracker gimballed to a pair of star trackers. Typical numerical examples and accuracy figures are developed to show that the method yields accurate results.

Presented at the ARS 15th Annual Meeting, Washington, D. C., Dec. 5-8, 1960.

¹ Director of Research. Member ARS.

² Member of Research Staff.

WITHIN the next decade the satellite traffic, consisting of scientific and military satellites originating in the U. S. and abroad, will increase and saturate ground based systems used for satellite navigation. Also, manned satellites, which will certainly require on-board navigation systems, are not far from being a reality. Thus the development of self-contained satellite navigation systems appears inevitable.

The Navigation Problem

In its simplest form, the satellite navigation problem consists of two basic parts: *point dynamics*, defining the motion of the spacecraft, treated as a point mass, around the geocenter; and *rigid body dynamics*, pertaining to the sensing and establishment of the satellite attitude with reference to a previously selected reference coordinate system. The solution of these two problems is largely affected by practical considerations relating to the mission of the satellite, the equipment design considerations, and the time period during which the navigation system is desired to operate.

The referenced literature describes many self-contained satellite navigational schemes. Some of the promising schemes for satellite position and attitude sensing (1)³ are: *V/H*, velocity over height technique; differential star aberration; occultation of the stars by the moon; solar-lunar tracking; and star-Earth tracking. The selection of any of these schemes is often based on considerations such as the achievable accuracy; size, weight and power requirements; the complexity and degree of data processing; ability to withstand space environmental considerations; cost; producibility, and—perhaps the most important of all—reliability over a prolonged operating period, and overall system simplicity.

Suggested Navigation System Configuration

This paper is built around the concept of sensing the satellite attitude by means of a prime reference established by a pair of star trackers; direct measurements, over a period of time, of the local geocentric vertical by means of a horizon tracker; and determining the orbital elements by a precision time reference and a data processing computer.

The basic conceptual feature of this approach is that suitable stars are preselected and tracked. The initial star lock-on can be achieved by orienting the satellite from the injection guidance subsystem's gyroscopes. Once this lock-on is achieved, it is continually maintained to obviate a complex search and identification process implied by loss of orientation resulting from extended shutdowns of operation.

³ Numbers in parentheses indicate References at end of paper.

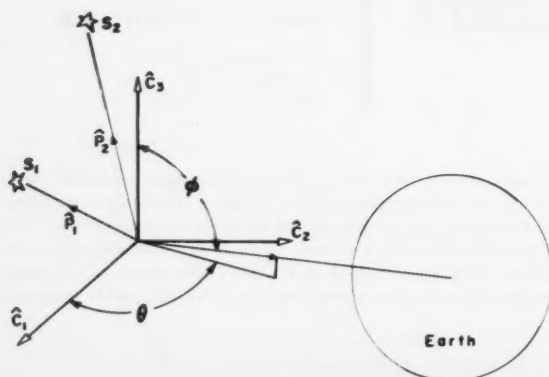


Fig. 1 Basic geometry of the suggested method

A typical design of a lightweight star tracker is described in (2). Two star trackers are locked on to two known stars S_1 and S_2 as shown in Fig. 1. The selection of the stars S_1 and S_2 is based on the techniques described in (3). A coordinate reference frame consisting of three mutually perpendicular vectors C_1 , C_2 , C_3 is rigidly connected to the two-star tracker telescopes, which are represented by the vectors P_1 and P_2 in Fig. 1. With the star trackers locked on to the stars S_1 and S_2 , the orientation of the frame (C_1 , C_2 , C_3) in inertial space is known. To avoid large errors in measurement of angle θ resulting from orientations in which angle ϕ is near zero it is assumed that this orientation is such that the vector C_3 is approximately normal to the plane of the orbit.

The azimuth and elevation angles θ and ϕ between the inertial frame C_1 , C_2 , C_3 and the local vertical are measured by means of a horizon tracker. A number of horizon trackers have been described in (2, 3, 4 and 5); such trackers permit the establishment of a vector between the satellite, as a point mass, and the center of the circular horizon as seen from the satellite.

Horizon trackers employing infrared sensors are generally more desirable than those using the visible part of the spectrum, since the former sense the entire circular part of the horizon due to the thermal considerations, whereas the latter devices require compensation for partial illumination or shadow effects.

The analysis that follows is applicable to any type of horizon tracker, so long as it establishes the local vertical based on observation of the local horizon. The horizon tracker may be equipped with gimbal pickoffs for measuring the angles θ and ϕ in Fig. 1.

Discussion of the time measurements and data processing computer is postponed until after a full description of the theory of the suggested method is given; such considerations will be developed in a subsequent paper.

Theory of the Suggested Self-Contained Navigation System

In the ideal case of an exact knowledge of the force field about Earth, one would be able to precisely determine the trajectory of an Earth satellite if exact values of the position and velocity of the satellite at some instant in time were known. In practice, of course, neither of these requirements is met. Approximate values of position and velocity are used as initial conditions for the integration of equations of motion derived from an approximate model of the force field. This process results in predicted values of position and velocity which contain errors that grow with time. The method of navigation presented in this paper reduces these errors by using a large number of measurements of the direction of the local vertical to obtain better estimates of the initial conditions and by limiting the time period over which the force field model is used to predict position and velocity.

The force model chosen for this paper is the Jeffreys model of the oblate Earth. This model neglects atmospheric drag and perturbations such as the influence of the moon, but these effects appear to be negligible for typical Earth satellite trajectories for periods of time corresponding to at least several cycles of the satellite. If the only requirement is that the model give a satisfactory representation of a satellite's trajectory for a couple of periods, the assumed model is probably adequate for the present purpose.

Having chosen the force field model the problem of trajectory determination reduces to that of determining the six initial conditions of position and velocity, henceforth referred to as the orbital elements.

Let us suppose that n measurements of the angles θ and ϕ are made using the basic geometry shown in Fig. 1. The data will then consist of the triples

$$(\theta_1, \phi_1, t_1), (\theta_2, \phi_2, t_2), \dots, (\theta_n, \phi_n, t_n)$$

where t_i is the time at which the i th measurement is taken. We now desire to use this data to determine which trajectory among all those possible under the chosen force field model conforms most closely to the direct measurements. Each of the angles θ and ϕ is a member of a six parameter family of curves in the time domain, that is

$$\theta = \theta(a_1, a_2, a_3, a_4, a_5, a_6, t) \quad [1]$$

$$\phi = \phi(a_1, a_2, a_3, a_4, a_5, a_6, t) \quad [2]$$

where a_1, a_2, a_3 , etc. denote the values of the orbital elements. The following section describes a method of calculating those values of the orbital elements which describe the curves that best fit the measured data.

Calculation of Orbital Elements

A criterion is needed for deciding when curves of the form given in Eqs. 1 and 2 fit the measured data as closely as possible. The criterion chosen is that the quantity L , defined by the equation

$$L(a_1, a_2, a_3, a_4, a_5, a_6) = \sum_{i=1}^n [\theta(a_1, a_2, a_3, a_4, a_5, a_6, t_i) - \theta_i]^2 + [\phi(a_1, a_2, a_3, a_4, a_5, a_6, t_i) - \phi_i]^2 \quad [3]$$

should be as small as possible where

- n = number of measurements
- θ_i = i th measurement of θ
- ϕ_i = i th measurement of ϕ
- t_i = time of i th measurement

$$L(a_1, \dots, a_6) = \sum_{i=1}^n \left\{ \left[\theta_c(t_i) + \sum_{j=1}^6 K_j(t_i)(a_j - a_j^*) \right] - \theta_i \right\}^2 + \sum_{i=1}^n \left\{ \left[\phi_c(t_i) + \sum_{j=1}^6 J_j(t_i)(a_j - a_j^*) \right] - \phi_i \right\}^2 \quad [11]$$

The trajectory of the satellite will be computed with respect to an Earth centered inertial coordinate system (C_1', C_2', C_3'), which has its coordinate axes parallel to the frame (C_1, C_2, C_3) described in Fig. 1. Let x, y, z denote the components of satellite position and $\dot{x}, \dot{y}, \dot{z}$ denote the components of satellite velocity in this Earth centered coordinate system.

Suppose that approximate values for position and velocity at time t_0 are known and let these values be used as approximate orbital elements. Denote these approximate orbital elements as $a_1^*, a_2^*, \dots, a_6^*$. As will be shown later, it is

$$\sum_{j=1}^6 \sum_{i=1}^n [K_k(t_i)K_j(t_i) + J_k(t_i)J_j(t_i)] \xi_j = \sum_{i=1}^n [\theta_i - \theta_c(t_i)] K_k(t_i) + [\phi_i - \phi_c(t_i)] J_k(t_i) \quad [13]$$

then possible to use the force equations resulting from the Jeffreys model to obtain values for

$$\theta(a_1^*, \dots, a_6^*; t_i)$$

$$\phi(a_1^*, \dots, a_6^*; t_i)$$

$$\left. \begin{aligned} \left[\frac{\partial}{\partial a_j} \theta(a_1, \dots, a_6; t_i) \right] a_k &= a_k^* \\ \left[\frac{\partial}{\partial a_j} \phi(a_1, \dots, a_6; t_i) \right] a_k &= a_k^* \end{aligned} \right\} \begin{aligned} j &= 1, 2, 3, 4, 5, 6 \\ k &= 1, 2, 3, 4, 5, 6 \end{aligned} \quad [4]$$

where t_i = time of i th measurement.

For convenience in notation we define the following quantities

$$\theta_c(t_i) \equiv \theta(a_1^*, \dots, a_6^*; t_i) \quad [5]$$

$$\phi_c(t_i) \equiv \phi(a_1^*, \dots, a_6^*; t_i) \quad [6]$$

$$K_j(t_i) \equiv \left[\frac{\partial}{\partial a_j} \theta(a_1, \dots, a_6; t_i) \right] a_k = a_k^* \quad [7]$$

$$J_j(t_i) \equiv \left[\frac{\partial}{\partial a_j} \phi(a_1, \dots, a_6; t_i) \right] a_k = a_k^* \quad [8]$$

$$j = 1, 2, 3, 4, 5, 6, k = 1, 2, 3, 4, 5, 6$$

If we assume that the approximate values of the initial conditions are close enough to the true values that we may neglect all terms higher than first order in a Taylor series expansion of θ and ϕ about $a_1^*, a_2^*, \dots, a_6^*$, it then follows

that

$$\theta(a_1, \dots, a_6; t_i) = \theta_c(t_i) + \sum_{j=1}^6 K_j(t_i)(a_j - a_j^*) \quad [9]$$

$$\phi(a_1, \dots, a_6; t_i) = \phi_c(t_i) + \sum_{j=1}^6 J_j(t_i)(a_j - a_j^*) \quad [10]$$

Substituting into Eq. 3 we obtain

It is necessary to compute the values of a_1, \dots, a_6 which minimize L . Defining the quantity ξ_j by

$$\xi_j = a_j - a_j^* \quad [12]$$

one can see that a necessary condition for L to be a minimum is

$$\frac{\partial L}{\partial a_k} = \frac{\partial L}{\partial \xi_k} = 0 \quad k = 1, 2, 3, 4, 5, 6$$

Performing the indicated differentiations and rearranging the following six linear equations in the six unknowns, ξ_j are obtained

for $k = 1, 2, 3, 4, 5, 6$. If we let

$$\bar{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \end{pmatrix} \quad [14]$$

and \bar{C} be the 6×6 matrix whose (k, j) element is

$$C_{kj} = \sum_{i=1}^n [K_k(t_i)K_j(t_i) + J_k(t_i)J_j(t_i)] \quad [15]$$

$$\bar{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_6 \end{pmatrix} \quad [16]$$

where

$$d_k = \sum_{i=1}^n [\theta_i - \theta_e(t_i)] K_k(t_i) + (\phi_i - \phi_e(t_i)) J_k(t_i) \quad [17]$$

for $k = 1, 2, 3, 4, 5, 6$. Then the equations can be expressed in matrix form as follows

$$\bar{C}\bar{\xi} = \bar{d} \quad [18]$$

$\bar{\xi}$ will then be given by the expressions

$$\bar{\xi} = \bar{C}^{-1}\bar{d} \quad [19]$$

where \bar{C}^{-1} is the inverse of the matrix \bar{C} . Let

$$\bar{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_6 \end{pmatrix} \quad \bar{a}^* = \begin{pmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_6^* \end{pmatrix} \quad [20]$$

It follows that

$$\bar{a} = \bar{a}^* + \bar{C}^{-1}\bar{d} \quad [21]$$

The matrix \bar{a} given by this equation will in general provide a better estimate of the orbital elements than the matrix \bar{a}^* does. If advisable, the value of \bar{a} would then be used in place of \bar{a}^* in the foregoing procedure to obtain an even better estimate.

The process of obtaining a better estimate of the orbital elements is reduced to the problem of obtaining the elements of the matrices \bar{C} and \bar{d} and then solving Eq. 21. Examination of Eqs. 15 and 18 shows that in order to obtain these matrix elements it is necessary to have values for $\theta_e(t_i)$, $\phi_e(t_i)$, $K_i(t_i)$ and $J_i(t_i)$. A method for obtaining these values follows.

Let q_1, q_2, q_3 be the direction cosines of Earth's axis with respect to the frame (C_1', C_2', C_3').

Let the following quantities be defined

$$r = \sqrt{x^2 + y^2 + z^2} \quad [22]$$

$$U = (q_1x + q_2y + q_3z)/r \quad [23]$$

Then the normalized force field derived from the Jeffreys oblate Earth model is given by the equation

$$\mathbf{F}(x, y, z) = F_1(x, y, z)\mathbf{C}_1' + F_2(x, y, z)\mathbf{C}_2' + F_3(x, y, z)\mathbf{C}_3' \quad [24]$$

where

$$F_1(x, y, z) = g R_e \left[\frac{-R_e x}{r^3} - \frac{J R_e^3}{r^4} \left(\frac{x}{r} - \frac{5U^2 x}{r} + 2q_1 U \right) - \frac{D}{7} \frac{R_e^5}{r^6} \left(63 \frac{U^4 x}{r} - 42 \frac{U^2 x}{r} + \frac{3x}{r} - 28q_1 U^3 + 12q_1 U \right) \right] \quad [25]$$

where

$$g = 32.145 \text{ ft/sec}^2$$

$$R_e = \text{equatorial radius of Earth, } 3441.69 \text{ n mi}$$

$$J = 1.637 \times 10^{-3}$$

$$D = 10.6 \times 10^{-6}$$

$$\dots\dots\dots [26]$$

Similar expressions hold for F_2 and F_3 . From these expressions it may be seen that the equations of motion of a

point mass acted upon by this force field are

$$\begin{aligned} \ddot{x} &= F_1(x, y, z) \\ \ddot{y} &= F_2(x, y, z) \\ \ddot{z} &= F_3(x, y, z) \\ &\dots\dots\dots [27] \end{aligned}$$

Using the initial conditions a_1^*, \dots, a_6^* these equations may be integrated numerically to obtain values for $x, y, z, \dot{x}, \dot{y}, \dot{z}$ at any desired time. Since the values obtained will depend not only upon the time, but upon the initial conditions used, we may express the position and velocity as functions of the orbital elements and time, that is

$$x = x(a_1, \dots, a_6; t) \quad [28]$$

with similar expressions for the other position and velocity coordinates. In particular, the value for one of the coordinates, say x , found by integrating to time t , using initial conditions a_1^*, \dots, a_6^* would be

$$x = x^*(a_1^*, \dots, a_6^*; t_i) \quad [29]$$

which value we define as $x^*(t_i)$ with $y^*(t_i)$, etc., defined in a similar manner.

Examination of Fig. 1 shows that the following relationships must hold

$$\begin{aligned} \cos \theta_e &= -y/\sqrt{x^2 + y^2} \\ \sin \theta_e &= -x/\sqrt{x^2 + y^2} \\ \cos \theta_e &= -z/\sqrt{x^2 + y^2 + z^2} \\ &\dots\dots\dots [30] \end{aligned}$$

By substituting of the values x^*, y^* and z^* into Eq. 30 and using fixed table look-up or other methods of obtaining inverse trigonometric function values we may obtain $\theta_e(t_i)$ and $\phi_e(t_i)$.

It should be noted in the foregoing that the values obtained for θ_e and ϕ_e are the values at times corresponding to times of observation of the angles θ_i and ϕ_i . It is now necessary to obtain values for the partial derivatives of θ_e and ϕ_e at the times of observation.

Noting again that the values of the position and velocity coordinates obtained by the integration process that has been indicated are functions of the initial conditions and time, let us define the following 36 quantities

$$\begin{aligned} f_{1j}(t) &= \left[\frac{\partial}{\partial a_j} x(a_1, \dots, a_6, t) \right]_{a_j = a_j^*} \\ f_{2j}(t) &= \left[\frac{\partial}{\partial a_j} y(a_1, \dots, a_6, t) \right]_{a_j = a_j^*} \\ &\dots\dots\dots \\ f_{6j}(t) &= \left[\frac{\partial}{\partial a_j} z(a_1, \dots, a_6, t) \right]_{a_j = a_j^*} \\ &\dots\dots\dots [31] \end{aligned}$$

for $j = 1, 2, 3, 4, 5, 6$. Then the following system of 36 differential equations must hold

$$\begin{aligned} (d/dt)f_{1j}(t) &= f_{4j}(t) \\ (d/dt)f_{2j}(t) &= f_{5j}(t) \\ (d/dt)f_{3j}(t) &= f_{6j}(t) \\ \dots\dots\dots \\ \frac{d}{dt}f_{4j}(t) &= \frac{\partial F_1}{\partial x} f_{1j}(t) + \frac{\partial F_1}{\partial y} f_{2j}(t) + \frac{\partial F_1}{\partial z} f_{3j}(t) \end{aligned}$$

$$\begin{aligned} \frac{d}{dt} f_{1j}(t) &= \frac{\partial F_2}{\partial x} f_{1j}(t) + \frac{\partial F_2}{\partial y} f_{2j}(t) + \frac{\partial F_2}{\partial z} f_{3j}(t) \\ \frac{d}{dt} f_{2j}(t) &= \frac{\partial F_3}{\partial x} f_{1j}(t) + \frac{\partial F_3}{\partial y} f_{2j}(t) + \frac{\partial F_3}{\partial z} f_{3j}(t) \\ &\dots\dots\dots [32] \end{aligned}$$

for $j = 1, 2, 3, 4, 5, 6$. From Eq. 25 explicit expressions may be obtained for the factors $\partial F_i/\partial x, \partial F_i/\partial y, \partial F_i/\partial z, i = 1, 2, 3$. Then using the initial conditions

$$\begin{aligned} [f_{ij}]_{t=t_0} &= 1 & i &= j \\ [f_{ij}]_{t=t_0} &= 0 & i &\neq j \end{aligned} \dots\dots\dots [33]$$

Eqs. 32 may be integrated numerically to time t_i and values obtained for

$$\begin{aligned} f_{1j}(t_i) \\ f_{2j}(t_i) \\ f_{3j}(t_i) \\ j = 1, 2, 3, 4, 5, 6 \end{aligned} \dots\dots\dots [34]$$

By differentiating Eqs. 30 the following relationships may be obtained

$$\begin{aligned} \frac{\partial \theta_c}{\partial a_j} &= \frac{y}{x^2 + y^2} f_{1j} - \frac{x}{x^2 + y^2} f_{2j} \\ \frac{\partial \phi_c}{\partial a_j} &= \frac{1}{x^2 + y^2 + z^2} \times \\ &\left(\frac{-zx}{\sqrt{x^2 + y^2}} f_{1j} - \frac{zy}{\sqrt{x^2 + y^2}} f_{2j} + \sqrt{x^2 + y^2} f_{3j} \right) \end{aligned} \dots\dots\dots [35]$$

for $j = 1, 2, 3, 4, 5, 6$. By substituting the values for x^*, y^*, z^* and $f_{1j}(t_i), f_{2j}(t_i), f_{3j}(t_i)$ we obtain

$$\begin{aligned} K_j(t_i) &= \left[\frac{\partial \theta_c}{\partial a_j} \right]_{t=t_i} \\ J_j(t_i) &= \left[\frac{\partial \phi_c}{\partial a_j} \right]_{t=t_i} \end{aligned} \dots\dots\dots [36]$$

and therefore are able to construct the matrix elements of \bar{C} and \bar{d} for time t_i .

System Flow Diagram

The process of updating the orbital elements can be described as follows. At the beginning, let the approximate orbital elements be a_1^*, \dots, a_6^* . These elements can be obtained by any of the following methods.

Use of injection guidance system, especially if the system is to begin computing immediately after injection into orbit.

Crude establishment of range and angle, and hence the orbital elements, by horizon tracker angular subtense measurement.

Measurements of angles between C_1, C_2 and C_3 and two known satellites.

RF tracking and measurement from another platform, such as Earth, and transmission of orbital data by a communication link.

Detailed discussions of the first and last of the forementioned are outside the scope of this paper. The second and third methods are of considerable interest and will be the subject of another paper.

Let time $t = \tau$ be the time for which the approximate orbital elements are known and let the period of updating or data smoothing be T . For times between $\tau + mT$ and $\tau + (m + 1)T$, where m is some integer, the system is in the state shown in Fig. 2. At $t = \tau + (m + 1)T$

$$\begin{aligned} a_1^*(\tau + mT) & \quad a_2^*(\tau + mT) & a_3^*(\tau + mT) \\ a_4^*(\tau + mT) & \quad a_5^*(\tau + mT) & a_6^*(\tau + mT) \end{aligned}$$

are erased and replaced by

$$\begin{aligned} a_1^*(\tau + (m + 1)T) & \quad a_2^*(\tau + (m + 1)T) & a_3^*(\tau + (m + 1)T) \\ a_4^*(\tau + (m + 1)T) & \quad a_5^*(\tau + (m + 1)T) & a_6^*(\tau + (m + 1)T) \end{aligned}$$

The data in storage are erased and new data are taken in during the time interval

$$\tau + (m + 1)T < t < \tau + (m + 2)T$$

In an unmanned satellite, the computer equipment and its design presents a very difficult problem. The approach to computer design must be to pursue scrupulously a data processing simplification program, and when this simplification is at or near the envisaged minimum, an extremely reliable computer hardware design must be developed.

If the required data processing problem is so complex that reliability considerations forbid a completely self-contained satellite computer, then a compromise solution might be to mechanize input and output equipment in the satellite equipped with a telemetry system to the ground, where the primary data handling would be done by computers which have the advantage of regular maintenance. This may be called a "partially self-contained system" and, of course, defeats some of the ground rules on which the original objectives of completely self-contained navigation were established.

Accuracy of Orbit Determination

The accuracy with which the star tracker geometry is mechanized is equivalent to a bias error in the reference coordinate system. This uncertainty can be reasonably specified as a value under 10 arc sec and, with special care, made as small as 2 arc sec.

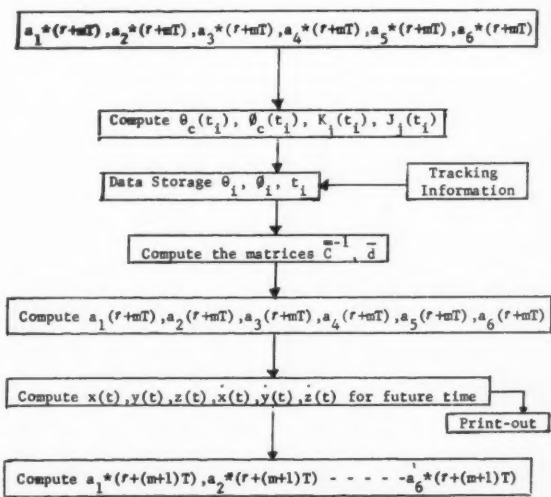


Fig. 2 Flow diagram of updating orbital elements

The errors in the horizon tracker measurements can be divided into two groups, each of which is subdivisible as follows.

Errors independent of intrinsic horizon tracker accuracy.

Irregularities of Earth's surface.

Oblateness.

Refraction of the radiation by the atmosphere.

Horizon tracker sensing of the horizon discontinuity at some height above the actual surface, due to atmospheric effects. This is to be distinguished from the errors arising in the location of this discontinuity by the horizon tracker.

Errors intrinsic to the horizon tracker.

Errors due to the uncertainty in the horizon tracker location of the horizon discontinuity; and errors in the angular references (scanning systems). These two items are to be associated with particular trackers and their sensor spectral response characteristics and other parameters which are well defined.

Errors in the electronic conversion of the sensed horizon into angular or range data.

Reference (6) describes detailed examination of the horizon tracker errors. The statistical analysis of a few cases suggests that horizon tracker accuracies of 6 arc min or 0.1 deg are very reasonable.

The angle pickoffs which relate the horizon tracker pointing direction to the reference coordinate system established by the stars are realistically obtainable in the form of multiple resolvers having fairly reasonable performance factors, with resolutions of less than 5 arc sec.

The cumulative effects of these prime measurements are errors in the angle measurements of Fig. 1.

Let θ_i' and ϕ_i' be the actual values of θ and ϕ at time t . Then

$$\begin{aligned}\theta_i' &= \theta_i + \epsilon_i \\ \phi_i' &= \phi_i + \delta_i \\ &\dots\dots\dots [37]\end{aligned}$$

where ϵ_i and δ_i represent the errors in direct angle measurements.

To obtain an idea of the accuracy of the forementioned method of orbit determination, it is assumed that $\epsilon_1, \epsilon_2, \dots, \epsilon_n, \delta_1, \delta_2, \dots, \delta_n$ are a random sample from a normal population with mean 0 and standard deviation σ .

Let $a_1', a_2', a_3', a_4', a_5', a_6'$ denote the actual values of the orbital elements. Let

$$d_k' = \sum_{i=1}^n [(\theta_i' - \theta_i(t))K_k(t_i) + (\phi_i' - \phi_i(t_i))J_k(t_i)] \quad [38]$$

Let

$$\bar{a}' = \begin{pmatrix} a_1' \\ a_2' \\ \vdots \\ a_6' \end{pmatrix} \quad \bar{d}' = \begin{pmatrix} d_1' \\ d_2' \\ \vdots \\ d_6' \end{pmatrix} \quad [39]$$

Then $\bar{d}' = \bar{a}' + C^{-1}\bar{d}$. In the orbital guidance system, one calculates \bar{a} from the equation

$$\bar{a} = \bar{a}' + \bar{C}^{-1}\bar{d} \quad [40]$$

Therefore

$$\bar{a}' - \bar{a} = \bar{C}^{-1}(\bar{d}' - \bar{d}) \quad [41]$$

Define

$$\Delta a_j = a_j' - a_j \quad [42]$$

and denote \bar{C}^{-1} by $\bar{B} = [B_{jk}]$. One then has

$$\begin{aligned}\Delta a_j &= \sum_{k=1}^6 B_{jk} \sum_{i=1}^n [(\theta_i' - \theta_i)K_k(t_i) + (\phi_i' - \phi_i)J_k(t_i)] \\ &= \sum_{k=1}^6 B_{jk} \left(\sum_{i=1}^n \epsilon_i K_k(t_i) + \delta_i J_k(t_i) \right) \\ &= \sum_{i=1}^n \sum_{k=1}^6 (B_{jk} K_k(t_i) \epsilon_i + B_{jk} J_k(t_i) \delta_i) \\ &\dots\dots\dots [43]\end{aligned}$$

Since it is assumed that $\epsilon_1, \epsilon_2, \dots, \epsilon_n, \delta_1, \delta_2, \dots, \delta_n$ are random samples from a normal population with mean 0 and standard deviation σ , it follows from a theorem of statistics that Δa_j will be normally distributed with mean 0 and standard deviation given by

$$\sigma(\Delta a_j) = \sigma \times \sqrt{\sum_{i=1}^n \left[\left(\sum_{k=1}^6 B_{jk} K_k(t_i) \right)^2 + \left(\sum_{k=1}^6 B_{jk} J_k(t_i) \right)^2 \right]} \quad [44]$$

Therefore by squaring the sums as indicated and grouping terms appropriately we obtain

$$\sigma(\Delta a_j) = \sigma \times \sqrt{\sum_{i=1}^n \sum_{k=1}^6 \sum_{l=1}^6 B_{jk} B_{jl} (K_k(t_i) K_l(t_i) + J_k(t_i) J_l(t_i))} \quad [45]$$

But

$$C_{kl} = \sum_{i=1}^n (K_k(t_i) K_l(t_i) + J_k(t_i) J_l(t_i)) \quad [46]$$

Therefore

$$\sigma(\Delta a_j) = \sigma \sqrt{\sum_{k=1}^6 \sum_{l=1}^6 B_{jk} B_{jl} C_{kl}} \quad [47]$$

$$\sigma(\Delta a_j) = \sigma \sqrt{\sum_{l=1}^6 B_{jl} \sum_{k=1}^6 B_{jk} C_{kl}} \quad [48]$$

But $\bar{B} = \bar{C}^{-1}$, therefore $\bar{B}\bar{C} = [1]$, the unit matrix, and

$$\sum_{k=1}^6 B_{jk} C_{kl} = \begin{cases} 1 & \text{when } j = l \\ 0 & \text{when } j \neq l \end{cases} \quad [49]$$

Therefore

$$\sigma(\Delta a_j) = \sigma \sqrt{B_{jj}} \quad [50]$$

That is, the standard deviation of the error in the calculation of the orbital element a_j is just the square root of the j th diagonal term of the matrix \bar{C}^{-1} times the standard deviation of an individual measurement.

One now proceeds to calculate some numerical values for $\sigma(a_1), \sigma(a_2), \sigma(a_3), \sigma(a_4), \sigma(a_5), \sigma(a_6)$. In carrying out this program, it is found helpful to assume that the only force experienced by the satellite is due to the inverse square field of a spherical Earth, and that the satellite is in a near circular orbit. Suppose that

$$\begin{aligned}a_1^* &= x_0^* = -r^* \\ a_2^* &= y_0^* = 0 \\ a_3^* &= z_0^* = 0 \\ a_4^* &= \dot{x}_0^* = 0 \\ a_5^* &= \dot{y}_0^* = -V^* = -R_e \sqrt{g}/\sqrt{r^*} \\ a_6^* &= \dot{z}_0^* = 0 \\ &\dots\dots\dots [51]\end{aligned}$$

That is, the approximate knowledge of the orbital elements corresponds to a circular orbit of radius r^* and velocity V^* .

Suppose also that $t_1 = 0$, (i.e., time is measured from the taking of the first measurement). Let

$$\omega = \frac{R_e}{r^*} \sqrt{\frac{g}{r^*}} \quad [52]$$

Then, under the assumption of an inverse square field only, one has

$$\begin{aligned} K_1(t) &= (3\omega t - 2 \sin \omega t)/r^* \\ K_2(t) &= (1 - 2 \cos \omega t)/r^* \\ K_3(t) &= 0 \\ K_4(t) &= (2 - 2 \cos \omega t)/V^* \\ K_5(t) &= (3\omega t - 4 \sin \omega t)/V^* \\ K_6(t) &= 0 \\ J_1(t) &= 0 \\ J_2(t) &= 0 \\ J_3(t) &= \cos \omega t/r^* \\ J_4(t) &= 0 \\ J_5(t) &= 0 \\ J_6(t) &= \sin \omega t/V^* \end{aligned} \quad [53]$$

It is assumed that n measurements are made at equal time intervals over a single cycle.

Let $T^* = 2\pi/R_e \sqrt{g/r^*}$ be the orbital period of the circular reference orbit.

Let Δt be the time spacing between individual readings. Then

$$\Delta t = T^*/n \quad [54]$$

The kj th element of the matrix \bar{C} is given by

$$\begin{aligned} C_{kj} &= \sum_{i=1}^n (K_k(t_i) K_j(t_i) + J_k(t_i) J_j(t_i)) \\ &= \frac{n}{T^*} \sum_{i=1}^n (K_k(t_i) K_j(t_i) + J_k(t_i) J_j(t_i)) \Delta t \end{aligned} \quad [55]$$

It is assumed that n is large enough so that the last sum can be approximated by the integral

$$\int_0^{T^*} (K_k(t) K_j(t) + J_k(t) J_j(t)) dt$$

One then has

$$C_{kj} = \frac{n}{T^*} \int_0^{T^*} (K_k(t) K_j(t) + J_k(t) J_j(t)) dt \quad [56]$$

Therefore

$$\bar{C} = n \begin{bmatrix} \frac{14 + 12\pi^2}{r^{*2}} & \frac{3\pi}{r^{*2}} & 0 & \frac{6\pi}{r^* V^*} & \frac{22 + 12\pi^2}{r^* V^*} & 0 \\ \frac{3\pi}{r^{*2}} & \frac{3\pi}{r^{*2}} & 0 & \frac{4}{r^* V^*} & \frac{3\pi}{r^* V^*} & 0 \\ 0 & 0 & \frac{1}{2r^{*2}} & 0 & 0 & 0 \\ \frac{6\pi}{r^* V^*} & \frac{4}{r^* V^*} & 0 & \frac{6\pi}{V^{*2}} & \frac{6\pi}{V^{*2}} & 0 \\ \frac{22 + 12\pi^2}{r^* V^*} & \frac{3\pi}{r^* V^*} & 0 & \frac{6\pi}{V^{*2}} & \frac{32 + 12\pi^2}{V^{*2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2V^{*2}} \end{bmatrix} \quad [57]$$

\bar{C}^{-1} will then be the following matrix

$$\bar{C}^{-1} = \frac{r^{*2} V^{*2}}{2n(3\pi^2 - 18)} \begin{bmatrix} \frac{3\pi^2 + 32}{V^{*2}} & \frac{30\pi}{V^{*2}} & 0 & \frac{-30\pi}{r^* V^*} & \frac{-3\pi^2 - 22}{r^* V^*} & 0 \\ \frac{30\pi}{V^{*2}} & \frac{36\pi^2 - 108}{V^{*2}} & 0 & \frac{-30\pi^2 + 72}{r^* V^*} & \frac{-24\pi}{r^* V^*} & 0 \\ 0 & 0 & \frac{12\pi^2 - 72}{V^{*2}} & 0 & 0 & 0 \\ \frac{-30\pi}{r^* V^*} & \frac{-30\pi^2 + 72}{r^* V^*} & 0 & \frac{27\pi^2 - 54}{r^{*2}} & \frac{24\pi}{r^{*2}} & 0 \\ \frac{-3\pi^2 - 22}{r^* V^*} & \frac{-24\pi}{r^* V^*} & 0 & \frac{24\pi}{r^{*2}} & \frac{3\pi^2 + 14}{r^{*2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{12\pi^2 - 72}{r^{*2}} \end{bmatrix} \quad [58]$$

Thus, one obtains

$$\begin{aligned}\sigma(\Delta a_1) &= \sqrt{\frac{(3\pi^2 + 32)}{2(3\pi^2 - 18)}} \frac{r^*}{\sqrt{n}} \sigma \\ \sigma(\Delta a_2) &= \sqrt{\frac{(36\pi^2 - 108)}{2(3\pi^2 - 18)}} \frac{r^*}{\sqrt{n}} \sigma \\ \sigma(\Delta a_3) &= \sqrt{\frac{(12\pi^2 - 72)}{2(3\pi^2 - 18)}} \frac{r^*}{\sqrt{n}} \sigma \\ \sigma(\Delta a_4) &= \sqrt{\frac{(27\pi^2 - 54)}{2(3\pi^2 - 18)}} \frac{V^*}{\sqrt{n}} \sigma \\ \sigma(\Delta a_5) &= \sqrt{\frac{(3\pi^2 - 14)}{2(3\pi^2 - 18)}} \frac{V^*}{\sqrt{n}} \sigma \\ \sigma(\Delta a_6) &= \sqrt{\frac{(12\pi^2 - 72)}{2(3\pi^2 - 18)}} \frac{V^*}{\sqrt{n}} \sigma\end{aligned}\quad [59]$$

Simplifying these expressions, one obtains

$$\begin{aligned}\sigma(\Delta a_1) &= \frac{1.63}{\sqrt{n}} r^* \sigma \\ \sigma(\Delta a_2) &= \frac{3.26}{\sqrt{n}} r^* \sigma \\ \sigma(\Delta a_3) &= \frac{1.41}{\sqrt{n}} r^* \sigma \\ \sigma(\Delta a_4) &= \frac{3.03}{\sqrt{n}} V^* \sigma \\ \sigma(\Delta a_5) &= \frac{1.37}{\sqrt{n}} V^* \sigma \\ \sigma(\Delta a_6) &= \frac{1.41}{\sqrt{n}} V^* \sigma\end{aligned}\quad [60]$$

To get an idea of the accuracy with which the orbital ele-

ments can be determined, take

$$\begin{aligned}\sigma &= 1 \text{ min} = 1/3430 \text{ radians} \\ r^* &= 4900 \text{ miles} \\ V^* &= 23,400 \text{ fps} \\ n &= 1000\end{aligned}\quad [61]$$

One then has

$$\begin{aligned}\sigma(\Delta a_1) &= 420 \text{ ft} \\ \sigma(\Delta a_2) &= 840 \text{ ft} \\ \sigma(\Delta a_3) &= 370 \text{ ft} \\ \sigma(\Delta a_4) &= 0.71 \text{ fps} \\ \sigma(\Delta a_5) &= 0.32 \text{ fps} \\ \sigma(\Delta a_6) &= 0.33 \text{ fps}\end{aligned}\quad [62]$$

Conclusions

The mathematical theory has been established for the updating of orbital elements of a satellite by observing the local vertical with a horizon tracker gimballed to a primary inertial reference frame of coordinate axes established by a pair of star trackers in the satellite. A simple error analysis indicates that the smoothed values of the orbital elements are promisingly accurate.

Detailed considerations of system design and mechanization of the data processing functions, as well as comparison of the overall system with other self-contained navigation systems, are expected to be covered in subsequent papers.

Acknowledgment

The authors wish to acknowledge the assistance given by W. J. Reilly, member of the Research Staff at Nortronics, in the preparation of this paper.

References

- 1 Button, P. A., Mallory, P. E. and Boor, S. B., "V/H Satellite Attitude Control," Proc. IAS National Specialists Meeting on Guidance of Aerospace Vehicles, Boston, Mass., May 1960, p. 114.
- 2 Stevens, F., "Space Navigation by Star and Planet Tracking Techniques," Proc. IRE 5th Annual National Space Electronics and Telemetry Symposium, Washington, D. C., Sept. 1960, section 10.3.
- 3 Stevens, F., "Application of Optical Techniques to Interplanetary Navigation," Third AFOSR Symposium, Los Angeles, Calif., Oct. 14, 1960.
- 4 Arck, M. H. and Merlen, M. M., "Horizon Sensors for Vertical Stabilization of Satellites and Space Vehicles," Proc. IAS National Specialists Meeting on Guidance of Aerospace Vehicles, Boston, Mass., May 1960.
- 5 Kendall and Stalcup, "Attitude Reference Devices for Space Vehicles," presented at National Aeronautical Electronics Conf., Dayton, Ohio, May, 1959, Proc. IRE, April 1960.
- 6 Bishop, D. R., "Considerations of the Errors Involved in Horizon Tracker Determinations," Nortronics Research Memo. 2350-DRB-275 (under preparation).

